Drug Discovery and Development Technology Center
Division of Pharmaceutical Biology
Faculty of Pharmacy
University of Helsinki

Drug Discovery Screening and the Application of Genomics and Proteomics in the Drug Development Process for *Chlamydia pneumoniae*

Joni Alvesalo

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Pharmacy of the University of Helsinki,
for public criticism in Lecture Room 2 at Viikki infocenter (Viikinkaari 11)
on May 19th, 2007, at 12 o'clock noon.

HELSINKI 2007

Supervisor: Prof. Pia Vuorela, Ph.D.
Department of Biochemistry and Pharmacy
Åbo Akademi University
Finland

Reviewers: Professor Matthias Maass, MD, PhD
Institute of Medical Microbiology, Hygiene and Infectious Diseases
SALK Laboratory
Austria

Dr. Mirja Puolakkainen MD, PhD
Department of Viral Diseases and Immunology
National Public Health Institute
Finland

Opponent: Professor Riitta Lahesmaa, MD, PhD
Molecular Immunology Group
Turku Centre for Biotechnology
Finland

16/2007
ISBN 978-952-10-3909-6 (paperback)
ISBN 978-952-10-3910-2 (PDF, http://ethesis.helsinki.fi/)
ISSN 1795-7079

Yliopistopaino
Helsinki 2007

CONTENTS

1 ABSTRACT 6

2 LIST OF ORIGINAL PUBLICATIONS 8

3 LIST OF ABBREVIATIONS 9

4 INTRODUCTION 10

5 REVIEW OF THE LITERATURE 12

5.1 *Chlamydia pneumoniae* 12

5.1.1 History 12

5.1.2 Development cycle 12

5.1.3 Structure 14

5.1.4 Genome 15

5.1.5 Interaction with the host cell 15

5.1.6 Clinical manifestations 16

5.1.7 Treatment 17

5.1.8 Diagnostics 18

5.2 Drug discovery 19

5.2.1 Drug screening 19

5.2.1.1 Natural products 20

5.2.1.2 Synthetic libraries and combinatorial chemistry 21

5.2.1.3 Virtual screening 21

5.2.2 Genomics 22

5.2.2.1 Microarrays 22

5.2.3 Proteomics 23

5.2.3.1 Two-dimensional polyacrylamide gel electrophoresis 24

6 AIMS OF THE STUDY 25

7 EXPERIMENTAL 26

7.1 Materials 26

7.1.1 Cell line and chlamydia strains (**I -V**) 26

7.1.2 Pure compounds and synthetic derivatives (**I**) 26

7.1.3 Antimicrobial agents (**II**) 28

7.1.4 Synthetic compounds (**III**) 28

7.1.5 Microarrays and scanner (**IV**) 29

7.2 Methods 30

7.2.1 Mycoplasma detection (**I, IV, V**) 30

7.2.2 Infection protocol (**I - V**) 30

7.2.3 Cytotoxicity testing (**I, III, IV**) 31

7.2.4 Nucleic acid isolation and q-PCR (**II, IV**) 31

7.2.5 *C. pneumoniae* detection (**I – IV**) 31

7.2.6 Virtual screening (**III**) 32

7.2.7 Gene silencing **(IV)** 32

8 RESULTS AND DISCUSSION 33

8.1 Screening of active compounds (**I, II, III**) 33

8.1.1 Natural products (**I**) 34

8.1.2 Development of a new screening assay for *C. pneumoniae* (**II**) 37

8.1.3 Synthetic compounds (**III**) 39

8.2 Host cell response to *C. pneumoniae infection* (**IV, V**)....41

8.2.1 Interpretation of gene expression data (**IV**)....41

8.2.2 Selection of potential drug target genes (**IV**)....43

8.2.3 Drug target validation (**IV**)....44

8.2.4 Effect of target gene silencing on *C. pneumoniae* infection (**IV**)....45

8.2.5 Changes at protein level (**V**)....47

9 CONCLUSIONS....49

10 ACKNOWLEDGEMENTS....51

11 REFERENCES....53

1 ABSTRACT

Chlamydia pneumoniae is a common obligate intracellular bacterium that causes upper and lower respiratory infections worldwide. *C. pneumoniae* infections are often persistent, and an acute infection may sometimes turn chronic. The treatment of acute infections can be problematic. Prolonged high-dose treatment with the currently available antibiotics is often needed to achieve clinical cure. In addition, more serious diseases such as atherosclerosis have recently been associated with chronic *C. pneumoniae* infection, which has proven to be extremely difficult to diagnose and impossible to treat with current antibiotics. Thus, new antichlamydial compounds are urgently needed.

In the present study, new antichlamydial compounds were searched among natural and naturally derived compounds as well as from synthetic libraries. A library of structurally diverse natural compounds known to have a variety of health-promoting effects in humans was tested against *C. pneumoniae*. Also, a library of synthetic compounds was pre-screened *in silico* with a structural homolog of the potential *C. pneumoniae* target protein, in order to create a smaller and more active targeted library. Although antichlamydial compounds were found from both libraries, nature-based compounds showing higher activity.

A new high-throughput screening (HTS) assay was developed for a more efficient search of antichlamydial compounds. After optimization of this partly automated time-resolved fluorometric immunoassay (TR-FIA), we were able to obtain results that were reliable, reproducible, and consistent with the two most widely used assays in chlamydia susceptibility testing, immunofluorescence staining (IF staining) and quantitative polymerase chain reaction (q-PCR).

Although it has been known for a long time that *C. pneumonia* manipulates several host cell functions, details of these systems and the host cell signaling pathways related to this manipulation remain poorly characterized. Characterization of the host cell response to *C. pneumoniae* infection was one of the central purposes of this study. This response was monitored at both gene and protein levels. We were able to identify four different host cell proteins with significant differences in their expression caused by *C. pneumoniae* infection. All of these proteins were structural proteins, which supports earlier observations about the structural rearrangements that are required for a successful infection.
At the genome level, we were able to identify hundreds of genes that were significantly affected by infection. In order to get more information out of the microarray data, we used a gene ontology (GO) classification, which groups genes according to their biological processes, cellular components, and

molecular functions. One set of potential drug target genes assumed to be important for *C. pneumoniae* infection was selected using GO classification data, and another set was selected using gene-wise analysis data, where data from infected cells is compared against control cell data. Expression changes of all target genes were confirmed with q-PCR, validated genes were silenced with the corresponding small interfering RNA (siRNA) molecule, and the effect of silencing on *C. pneumoniae* infection was monitored. Both methods revealed genes whose silencing reduced the number of *C. pneumoniae* particles, but GO-classified data proved to be more accurate and much more informative compared to gene-wise analysis data.

2 LIST OF ORIGINAL PUBLICATIONS

I J. Alvesalo, H. Vuorela, P. Tammela, M. Leinonen, P. Saikku, P. Vuorela; Inhibitory effect of dietary phenolic compounds on *Chlamydia pneumoniae* in cell cultures. Biochemical Pharmacology 71 (2006) 735–741.

II P. Tammela, J. Alvesalo, L. Riihimäki, S. Airenne, M. Leinonen, P. Hurskainen, K. Enkvist, and P. Vuorela; A novel time-resolved fluorometric immunoassay for screening of antichlamydial compounds. Analytical Biochemistry 333 (2004) 39-48.

III J. Alvesalo, A. Siiskonen, M. Vainio, P. Tammela and P. Vuorela; Similarity based virtual screening: A tool for targeted library design. Journal of Medicinal Chemistry 49 (2006) 2353-2356.

IV J. Alvesalo, D. Greco, M. Leinonen, P. Vuorela and P. Auvinen; Microarray analysis of *Chlamydia pneumoniae* infected human epithelial cell line using gene ontology hierarchy (submitted).

V K. Savijoki, J. Alvesalo, P. Vuorela, M. Leinonen and N. Kalkkinen; Proteomic Profiling of the Human Lung Cell Response to *Chlamydia pneumoniae* K7 Infection Using Two-Dimensional Difference Gel Electrophoresis and Mass Spectrometry (submitted).

These publications are referred to in the text by their Roman numerals.

3 LIST OF ABBREVIATIONS

2D-DIGE	two-dimensional difference gel electrophoresis
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
DMSO	dimethyl sulfoxide
EB	elementary body
GO	gene ontology
HTS	high-throughput screening
IF staining	immunofluorescence staining
IFU	inclusion-forming unit
LPS	lipopolysaccharide
MIC	minimum inhibitory concentration
MIF	microimmunofluorescence
mRNA	messenger RNA
MS	mass spectrophotometer
NCE	new chemical entity
p.i.	post infection
PCR	polymerase chain reaction
q-PCR	quantitative polymerase chain reaction
RB	reticulate body
S/B	signal-to-background ratio
S/N	signal-to-noise ratio
siRNA	small interfering RNA
TR-FIA	time-resolved fluorometric immunoassay
Z’	screening window coefficient calculated from control sample data

4 INTRODUCTION

The genus *Chlamydia* was established in 1966, when Moulder reported the bacterial nature of this pathogen (Moulder, 1966), and *Chlamydia pneumoniae* was separated as a distinct species in 1992. *C. pneumoniae* is primarily a human pathogen. It is a very common bacterium worldwide, and almost everyone is infected at some point of their life. Serological studies indicate that 60 to 70% of the adult population have antibodies against it, men being more often seropositive than women (Grayston, 1992). An acute *C. pneumoniae* infection causes upper and lower respiratory track infections, which can be severe in elderly people. *C. pneumoniae* infections are estimated to cause about 10% of all community-acquired pneumonias and 5% of bronchitis and sinusitis cases in the adult population (Grayston, 2000).

In some cases, acute *C. pneumoniae* infection can become chronic. The molecular mechanisms leading to chronic infections are poorly understood, but *in vitro* chronic infection can be achieved with various methods (Hogan *et al.*, 2004). The *C. pneumoniae* life cycle provides ideal circumstances for the establishment of chronic infection. The elementary particle (EB), which is the infectious component of the bacterium, is metabolically inactive and therefore resistant to antibiotics. The intracellular form, i.e. the reticulate body (RB), is metabolically active and can be affected by some antibiotics, but eradication of the infection may be extremely difficult (Ekman *et al.*, 1993b), and *C. pneumoniae* can be recovered from cell cultures even after 30-day antibiotic treatment (Kutlin *et al.*, 2002). Over the last two decades, chronic *C. pneumoniae* infection has been associated with serious diseases such as atherosclerosis (Saikku *et al.*, 1988).

Acute *C. pneumoniae* infections can be treated with certain antibiotics (Kuo and Grayston, 1988) (Miyashita *et al.*, 1997) (Roblin and Hammerschlag, 1998), but symptoms may recur or the infection may become persistent despite of these treatments (Hammerschlag *et al.*, 1992b). And although there are problems concerning the treatment of acute *C. pneumoniae* infection, the situation with persistent infection is much worse. At the moment, there are no proper diagnostic methods or effective treatments for persistent *C. pneumoniae* infection. Incomplete eradication of persistent infection with the current antibiotics has been demonstrated *in vitro* (Kutlin *et al.*, 1999), and the same problem can also be seen in large-scale clinical trials (Andraws *et al.*, 2005) (Danesh, 2005).

Drug development is an extremely complex and risky business. It takes a long time and costs a lot of money, and there is no guarantee of success. Lately, the pharmaceutical industry has invested heavily in

high-throughput screening (HTS) platforms and robots. These investments have not paid off as anticipated (Schmid and Smith, 2004), since the majority of new drug candidates still fail in clinical trials, causing a decrease in the number of new chemical entities (NCE) (Graul and Prous, 2006). In order to improve the odds of success, a lot of effort has been put into the development of pre-clinical research methods. The availability of genome sequences for several pharmaceutically relevant organisms has opened up new possibilities for techniques that utilize this information, such as genomics and proteomics.

Genomics and proteomics refer to studies that are directed towards all genes and proteins of an organism, respectively. Using these techniques, it is possible to study both normal and pathological stages on a genome-wide scale and to acquire valuable information that can be utilized in the early phases of the drug development process. Implementation of these approaches in the drug discovery process will undoubtedly speed up the development of safer, more effective, and better targeted therapeutic agents (Onyango, 2004).

5 REVIEW OF THE LITERATURE

5.1 *Chlamydia pneumoniae*

5.1.1 History

The bacterium now known as *C. pneumoniae* was first isolated from the eye of a Taiwanese child back in 1965 and named TW-183. When the methods for cultivating chlamydia became available a few years later, it was noticed that TW-183 was similar, but not identical, to two already known chlamydia species, *Chlamydia psittaci* and *Chlamydia trachomatis*. Morphological comparisons indicated that TW-183 bears closer resemblance to *C. psittaci* than *C. trachomatis*. First suggestions about the role of TW-183 as human pathogen were discovered 1985 (Saikku *et al.*) and more specific information started to build up and after the first respiratory isolation (AR-39) (Grayston *et al.*, 1986). This isolation was done from a throat swab of a patient with pharyngitis in Seattle. It linked this pathogen, now named TWAR after the TW-183 and AR-39 isolates, to respiratory track infections and pneumonia. These findings led to a proposal for the establishment of a third species of the genus *Chlamydia*, called *Chlamydia pneumoniae* (Grayston *et al.,* 1989).

5.1.2 Development cycle

C. pneumoniae is a gram-negative obligatory intracellular parasite of the eukaryotic cell, which has a unique intracellular development cycle inside the inclusion in the host cell cytoplasm (**Fig. 1**). The development cycle of *C. pneumoniae* consists of two very different forms. The EB is a metabolically inactive form with a cystein-rich outer membrane. EB particles are about 0.3 μm in diameter, infective, and able to survive outside the host cell for a short time. The RB is larger, typically 1.0 μm in diameter, metabolically active, and able to replicate inside an inclusion in the host cell cytoplasm. *C. pneumoniae* infection begins when EB interact with the host cell, and the bacteria is taken up by the host cell via endocytosis. The details of this mechanism are unknown, but it is somehow related to the mannose 6-phosphate/insulin-like growth factor 2 receptor (Puolakkainen *et al.*, 2005). Inside the host cell, EBs are transported to the Golgi region, where they start to convert their endocytotic vesicle into an inclusion. Sphingomyelin and cholesterol are acquired from the host cell via an exocytic pathway, in a way that requires *C. pneumoniae* protein translation (Wolf and Hackstadt, 2001).

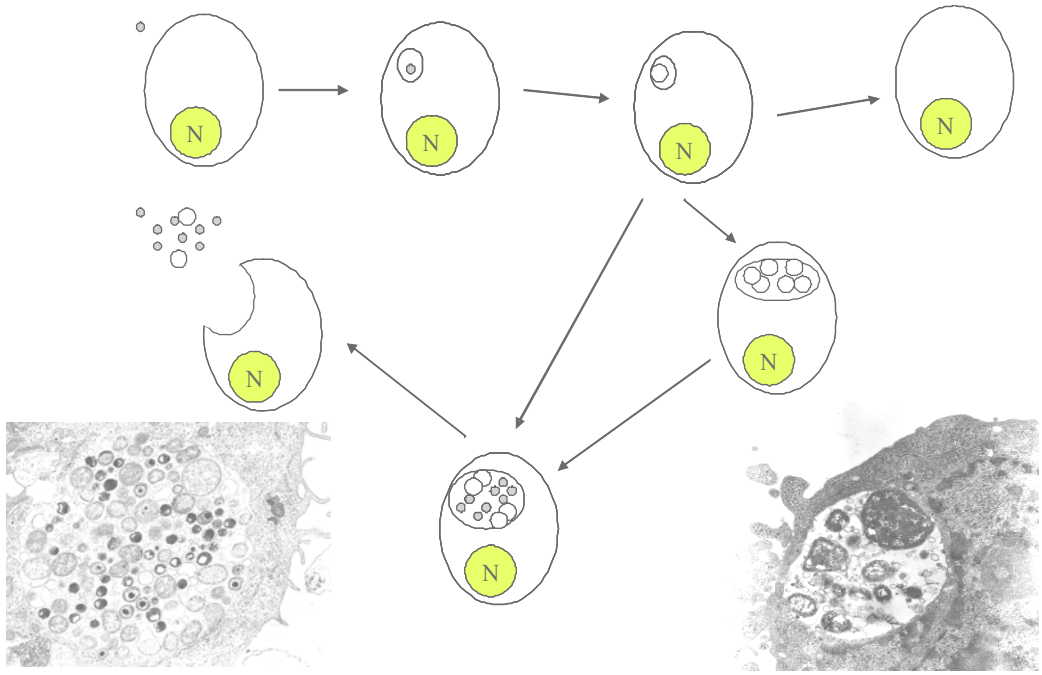

Figure 1. Developmental cycle of Chlamydia. Picture by Mirja Puolakkainen and electron microscopy by Laura Mannonen.

These molecules are used as construction material for the growing inclusion, which is a requirement for successful intracellular development cycle. After 8 h, EBs start to differentiate into RBs inside the inclusion, and about 19 h post-infection (p.i.) RBs start to replicate through binary fission (Wolf *et al.*, 2000). This continues up to 48 h p.i., after which the first RBs start to convert back to EBs, but intermediate forms as well as replicating RBs can still be found (**Fig. 2**). At the end of the development cycle at 72 h p.i., inclusions contain increasing amounts of pear-shaped EBs, which are characteristic for the *C. pneumoniae*, but also a considerable amount of RBs (Chi *et al.*, 1987). At about 90 h p.i., the intracellular cycle is complete, and the host cell has broken down in order to release the newly produced EBs into the extracellular space, where they can infect new cells.

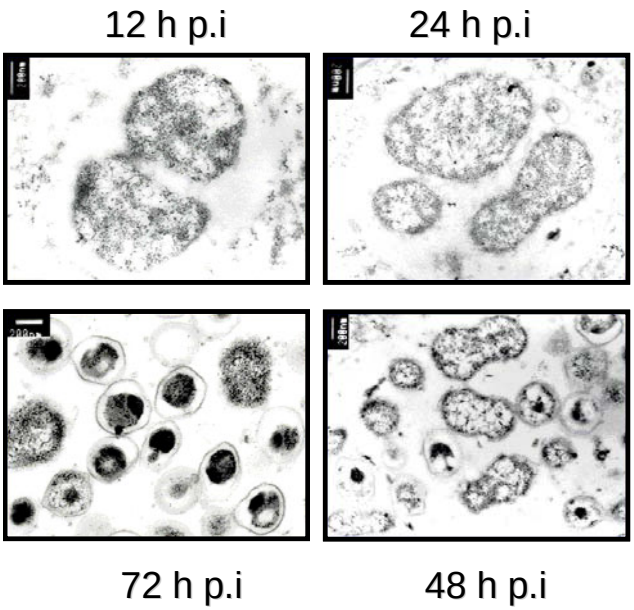

Figure 2. Electron microscopic images of the *C. pneumoniae* isolate K7 in HL cell culture at different stages of infection. Electron microscopy by Kari Lounatmaa.

5.1.3 Structure

The envelope of *C. pneumoniae* consists of two lipid layers, similarly to all gram-negative bacteria. Its outer membrane contains lipopolysaccharide (LPS), which is *Chlamydia*-specific and can be found in EBs and RBs. The inner layer contains peptidoglycan (Chopra *et al.*, 1998), which usually plays an important structural role in gram-negative bacteria, but in *Chlamydia* its amount is too small to participate significantly in the maintenance of the structural integrity of the membrane. Regardless of this, *Chlamydia* particles are quite strong. This rigidity and osmotic stability is thought to originate from the covalent sulphur bridges that are formed between cysteine-rich proteins in the outer membrane complex. This complex consist mostly of 40 kDa major outer membrane protein (MOMP) (Caldwell *et al.*, 1981), 60 kDa cysteine-rich outer membrane protein (omp2) (Newhall *et al.*, 1982), 12.5 kDa cysteine-rich outer membrane protein (omp3) (Hatch *et al.*, 1984), and several polymorphic membrane proteins (Pmp). Inside the envelope, *C. pneumoniae* has a condensed nucleoid structure, which can be visualized at the beginning of an infection (Chi *et al.*, 1987). This structure is a combination of DNA and histone-like proteins, which are not normally found in prokaryotes. The *Chlamydia* genome contains two genes for histone-like proteins, Hc1 and Hc2 (Kalman *et al.*, 1999). Hc2 has probably some role in

controlling stage-specific gene expression, whereas Hc1 is responsible for the silencing of gene expression at the end of the C*hlamydia* development cycle (Pedersen *et al.*, 1994).

5.1.4 Genome

The size of the *C. pneumoniae* genome is only about 1.2 million bases, and it contains 1072 likely protein-coding genes (Kalman *et al.*, 1999). This is a rather small genome even on the prokaryotic scale. The obvious reason for the small genome is the intracellular life style of *Chlamydia*, which enables it to acquire nutrients and structural components easily from the host cell. This reduces the amount of enzymes needed to synthesize various essential molecules. The lack of suitable methods for genetic manipulation of the *Chlamydia* genome has made it very difficult to study the functions of individual genes, but sequence comparisons with genes with known functions in other bacteria have proposed some role for 60% of the *C. pneumoniae* genes (Kalman *et al.*, 1999). According to these annotations, *C. pneumoniae* should have the tools to carry out independently some of the essential functions required for survival, but most of the essential pathways are incomplete and need certain molecules from the host cell. *Chlamydia* gene expression, which must be regulated accurately due to the very different needs and circumstances at the different stages of the development cycle, is controlled by the chlamydial sigma factors (Mathews and Timms, 2000).

5.1.5 Interaction with the host cell

C. pneumoniae is capable of multiplying in several different cell types (Kaukoranta-Tolvanen *et al.*, 1994). Inside the host cell, *C. pneumoniae* takes an active role and manipulates the host cell gene expression. This is accomplished partly via *inc* proteins (Rockey *et al.*, 1995), which are located on the cytoplasmic side of the inclusion membrane. Although the details of this interaction are unknown, it probably involves host cell kinases (Rockey *et al.*, 1997), which are responsible for the regulation of most host cell pathways. Another way to interact with the host cell is provided by chlamydial proteins, which are secreted into the host cell cytoplasm (Zhong *et al.*, 2001) (Shaw *et al.*, 2002) (Lugert *et al.*, 2004) (Vandahl *et al.*, 2005). The functions of these secreted proteins are mostly unknown, but at least some of them impair host cell defense mechanisms and possess proteolytic activity (Zhong *et al.*, 2001).

Microarray experiments with *C. pneumoniae* infected cells have yielded a lot of new information about the bacteria - host cell interaction. Coombes and Mahony (2001) studied the changes in 268 human

genes at various time points after *C. pneumoniae* infection (strain VR-1310) in a human microvascular endothelial cell line (HMCE-1) and found 20 genes with changed expression. This small number of affected genes encouraged them to conclude that the endothelial response to *C. pneumoniae* infection is quite limited at the messenger RNA (mRNA) level. Virok *et al.* (2003) did their microarray work with 2032 genes using a monocytic cell line (U937) and the *C. pneumoniae* strain AR-39 and found 128 genes that were affected by the infection. Most of the up-regulated genes they discovered were linked to pro-inflammatory functions, whereas most of the down-regulated genes were related to DNA and RNA metabolism or cell cycle regulation. Hess *et al.* (2003) used 12 000 genes in their microarray experiments and compared the changes in host cell (human cervical ephitelial HeLa cell) gene expression after *C. pneumoniae* infection to the changes following *C. trachomatis* and *Salmonella typhimurium* infections. They concluded that both chlamydia species has a highly similar, substantial, and actively controlled effect on the host cell gene expression. Mannonen *et al*. (2007) focused on the gene expression changes specific to persistent *C. pneumoniae* infection, induced by interferon gamma, and were able to establish several previously unknown modifications in host-cell gene expression. They used the HL cell line and found 29 genes (out of a total of ~4000) whose expression was changed significantly and over twofold due to the persistent infection. Shi and Tokunaga (2004) studied the effect of acute *C. pneumoniae* infection on umbilical vein endothelial cells, using a cDNA array with 588 human cardiovascular genes, and found support for the assumption that *C. pneumoniae* participates in atherosclerotic development *in vivo*. As a conclusion of these articles, it can be said that, although the authors found some common genes, the overall patterns of up- and down-regulated genes were very different. This is probably due to two factors: the biological aspect, meaning the different host cells, different *C. pneumoniae* strains, different amounts of *C. pneumoniae* in the infection protocols, etc., and the technical aspect, meaning the different microchips, different RNA labeling procedures, different data normalization, etc.

5.1.6 Clinical manifestations

C. pneumoniae spreads from person to person via respiratory droplets and can infect lung and respiratory cells. Although *C. pneumoniae* is a common bacterium, *C. pneumoniae* infections are diagnosed rarely, since most of them are asymptomatic or mild. In Finland, 10 – 20% of pneumonia cases (between epidemics) are caused by *C. pneumoniae* (Jokinen *et al.*, 2001). Acute *C. pneumoniae* infection can cause mild upper respiratory track infections, pharyngitis or sinusitis, and also lower respiratory track infections, such as bronchitis and pneumonia (Grayston *et al.*, 1990). The above-

mentioned disease are rarely life-threatening, but much more serious diseases have recently been associated with chronic *C. pneumonia* infection. This kind of infection can develop when the *Chlamydia* development cycle cannot be completed normally, and the bacterium converts to an aberrant, but viable persistent form (**Fig. 1**). The persistent form can be induced in different cell types by, for example, interferon-gamma (Beatty *et al.*, 1993), iron restriction (Raulston, 1997), or penicillin (Matsumoto and Manire, 1970), but it can also develop "naturally" without any additional substances in human monocytes (Koehler *et al.*, 1997). This natural way of persistence might be of extreme importance, since human immune cells (monocytes and macrophages) are capable of harboring the persistent form and spreading it around the body, at least in animal models (Moazed *et al.*, 1998). The most common and dangerous disease linked to chronic *C. pneumonia* infection is atherosclerosis. Saikku and colleagues discovered the connection between these two conditions in 1988 (Saikku *et al.*, 1988). After this, several studies have supported this finding (Epstein *et al.*, 1999), (Leinonen and Saikku, 2002), (Libby *et al.*, 1997), but the final causative evidence is still to come. Other severe diseases connected with chronic *C. pneumoniae* infection are asthma, chronic obstructive pulmonary disease (Hahn, 1999), and lung cancer (Kuo *et al.*, 1995), but as in the case of atherosclerosis, the role of *C. pneumoniae* in the pathogenesis of these diseases is still unclear.

5.1.7 Treatment

C. pneumoniae is susceptible to some antibiotics, such as erythromycin, tetracycline, and doxycycline (Kuo and Grayston, 1988), which are the most widely used treatments for acute infections, but some fluoroquinolones (Miyashita *et al.*, 1997) and ketolides (Roblin and Hammerschlag, 1998) have also demonstrated *in vitro* activity. Symptoms of *C. pneumonia* infection can recur, or infection can turn persistent despite treatment with these antibiotics (Hammerschlag *et al.*, 1992b), and as mentioned earlier, it is also possible to induce a chronic infection with certain antibiotics (Matsumoto and Manire, 1970). Therefore, more intensive long-term treatments are recommended (Kuo *et al.*, 1995), but there are no general guidelines for problematic situations, such as persistence or relapse of symptoms. In addition, more recent evidence suggests that *C. pneumoniae* infection in circulating human monocytes is refractory to antibiotic treatment (Gieffers *et al.*, 2001) (Baltch *et al.*, 2004). When these pieces of evidence are combined with the fact that there is no effective treatment for persistent *C. pneumoniae* infection, it is obvious that more research and new compounds are needed in this area.

5.1.8 Diagnostics

Acute *C. pneumoniae* infections are usually diagnosed using serology. A positive result requires a fourfold increase in IgM or IgG paired sera taken several weeks apart or single IgM $\geq$ 16 or single IgG $\geq$ 512 (Grayston *et al.*, 1990). Antibodies are generally measured with the microimmunofluorescence (MIF) test develop by Wang and Grayston (1970). This test uses EBs as antigens and is capable of measuring separately IgG and IgM antibody levels. This is important, since it helps to differentiate between past and recent infections as well as between primary and re-infections (Kuo *et al.*, 1995). Another way to study *C. pneumoniae* antibody levels is the complement fixation test, but this method has problems with specificity and sensitivity. On the positive side, it measures antibodies against chlamydial LPS, which are produced much earlier than IgA, IgG, or IgM antibodies, and it is also much easier to use than the MIF test.

Culturing has been most widely used method in *C. trachomatis* diagnostics, but these days it is replaced by different commercial nucleic acid amplification tests. Culturing is not suitable for *C. pneumoniae* diagnostics due to problems with isolation, successful growing, and sensitivity. Also, this technique requires good transportation and refrigeration facilities, in order to keep the bacteria infective.

Polymerase chain reaction (PCR) is capable of detecting small numbers of *C. pneumoniae* specific nucleic acids from a variety of samples. It can also be modified and used to determine whether *C. pneumoniae* is active or inactive. The PCR technique has superior sensitivity compared to the other methods, but it provides too many false negative and positive findings (Apfalter *et al.*, 2001). If these problems can be overcome, and standardized robust protocols can be developed, this method has the potential to take chlamydial diagnostics to the next level.

Although acute *C. pneumoniae* can sometimes be difficult to diagnose, there are much greater problems in the diagnostics of chronic infection, which is extremely difficult to distinguish from repeated infections by means of the current diagnostic methods. This is related to the fact that chronic *C. pneumoniae* infections can hide deep inside tissue where it is difficult to get decent samples. Usually, the diagnosis of chronic infection is based on IgG and especially IgA antibodies, but there are no clinical criteria for the diagnosis of chronic infection.

5.2 Drug discovery

Drug discovery is a multi-million dollar business. It has been estimated that pharmaceutical companies spend an average of 609 million euros to on the development of a new drug, of which sum about 40% is used in the pre-clinical phase (DiMasi *et al.*, 2003).
Bioactivity screening has been the central part of the drug development process for a long time. In this process, the activities of compounds are evaluated against a potential target molecule, disease model, or micro-organism. In the pharmaceutical industry, this screening is currently done in a high-throughput manner (DiMasi *et al.*, 2003). Compared to traditional screening systems, HTS is simple, rapid, effective, and economical. Automated HTS robots and workstations can screen over 100 000 compounds in a day (Liu *et al.*, 2004).

5.2.1 Drug screening

Advances in sequencing technology increased the number of potential drug targets dramatically in the early 1990s (Koehn and Carter, 2005). At the same time, the development of combinatorial chemistry techniques enabled the creation of enormous synthetic libraries, which could easily be screened in the HTS environment. Although natural products are excellent sources of drug candidates due to their high chemical diversity and biochemical specificity, the more difficult and time-consuming nature of natural compound libraries prompted pharmaceutical industry to use synthetic libraries (Newman *et al.*, 2003). This trend cost substantial amounts of money to pharmaceutical companies, but did not pay off in the expected way (Schmid and Smith, 2004), and the number of NCEs was declining at the beginning of the millennium (Graul and Prous, 2006). In retrospect, the biggest reason for this failure was probably the lack of structural diversity among synthetic compounds (Payne *et al.*, 2007). This view is also supported by several comparative studies on the structural diversity of synthetic and natural libraries (Feher and Schmidt, 2003), (Lee and Schneider, 2001), (Stahura *et al.*, 2000), which have agreed on the greater diversity of natural libraries.

Currently, pharmaceutical industry is in a situation where novel compounds are urgently needed. This is a problem, since it is not possible to create adequate diversity with the current synthetic methods, and since the natural products that possess this diversity are often unsuited to HTS (Feher and Schmidt, 2003). Pharmaceutical industry is addressing this problem in two ways: firstly, by making natural product libraries more suitable for HTS by, for example, using fractionated natural products libraries,

and secondly, by developing methods of combinatorial synthesis to enhance the structural complexity of synthesized compounds (Koehn and Carter, 2005).

5.2.1.1 Natural products

A natural product is a compound or substance produced by a living organism. Two groups of organisms, namely plants and microbes, have been the source and the innovation of most pharmaceutically relevant natural compounds. These biologically active compounds are usually secondary metabolites that are produced after external stimuli, such as infection (Strohl, 2000). Due to their specialized biochemical capabilities, plants and microbes can synthesize a vast array of structurally diverse chemical compounds, which can be used as protection against environmental factors. Between 1981 and 2002, 33% of new NCEs were purely synthetic. Of the remaining ones, 3% were vaccines, 5% were pure natural products, and the rest were somehow related to nature, e.g. nature-based pharmacophores, or derivatives of natural products (Newman *et al.*, 2003).

Microbes are the source of some of the best-selling drugs today, and they still continue to be the most promising source of new antibiotics (Pelaez, 2006). The revolutionary discovery of penicillin by the Scottish scientist Sir Alexander Fleming in 1928 was soon followed by the discovery of other antibiotics of bacterial or fungal origin, whose chemical scaffolds are still used in the development of new antibiotics. The major problem with these “old” antibiotics and their derivatives is the fast evolving resistance (Singh and Barrett, 2006). Although pharmaceutical industry is aware of this threat, there are economic, regulatory, and scientific reasons to explain the reluctance to develop new antibiotics (Nathan, 2004). As a consequence, only three new antibiotics classes have entered the marketd in the last 25 years (Butler and Buss, 2006).

Biologically active compounds have made several plants useful for man since ancient times as spices or medicines, and they still continue to be an inspiration and important source of new medicines for pharmaceutical industry. Plant-derived compounds are known to have a variety of beneficial effect on the human body. There are well-documented cases of the ability of plant compounds to control cell division and proliferation, platelet aggregation, detoxification, and inflammatory as well as immune responses (Manthey, 2000) (Middleton and Kandaswami, 1992), and there are also several plant-derived anti-cancer compounds in clinical use (Cragg and Newman, 2005). Despite the recent reduction in natural product operations in big pharmaceutical companies, medical plant discovery continues to

produce new lead molecules against various diseases such as Alzheimer's, malaria, and HIV/AIDS, but the full potential of plant-derived compounds can only be exploited if the above-mentioned challenges concerning HTS bioassays can be resolved (Balunas and Kinghorn, 2005).

5.2.1.2 Synthetic libraries and combinatorial chemistry

Synthetic libraries play an important role in the drug development process, but the basic problem with these libraries has traditionally been their lack of diversity (Payne *et al.*, 2007). Combinatorial chemistry has provided a partial solution of this problem. Combinatorial technology means synthetic technologies that are able to generate compound libraries rather than single compounds. Thousands of combinatorial libraries have been produced since the first small-molecule combinatorial library was introduced in the early 1990s (Gershell and Atkins, 2003). Originally, this technique was developed for the synthesis of peptide and oligonucleotide libraries (Furka *et al.*, 1991), but it was later modified to cover small molecules. The basic idea of combinatorial chemistry is to simultaneously synthesize large amounts of different compounds in a short period of time. Combinatorial techniques enable the combination of a number of chemical building blocks in different ways. In pharmaceutical industry, combinatorial chemistry is mainly used to create “focused” or “prospecting” libraries (Spaller *et al.*, 1997). Focused libraries consist of different modifications of a molecule with known activity and are used to find alternatives to this structure, whereas prospecting libraries consist of novel structures.

5.2.1.3 Virtual screening

HTS utilizes large compound libraries and sophisticated automated systems, which are often inaccessible to scientists outside the pharmaceutical industry. Virtual screening is a computer-assisted way to create smaller, focused libraries, against a known target. Such libraries can be created by using a target protein or the structure of a known ligand as a starting point. The former option requires a 3D crystal structure of the target protein. This 3D structure can also be a “comparative model”, which is created using the known 3D structures of closely related proteins. Small molecular structures are docked on the computer screen to the preferred place on the protein, in order to find the best binders. Ligand-based libraries use known active compounds as templates and try to create even more active compounds or find active compounds with better physicochemical properties. Since these focused libraries are rather small, they can be tested in a suitable bioassay without automation. They can also be designed for drug screening purposes by choosing only compounds with drug-like properties (Lipinski *et al.*, 1997). The

limitations of this technology are related to the accuracy of the different variables that specify compounds and targets, and when virtual compounds are used, to the number of compounds that can be actually synthesized based on their virtual model (Gershell and Atkins, 2003).

5.2.2 Genomics

Genomics means the study of an organism's entire genome. The interest for such research began to emerge more than 20 years ago, and it exploded after the first genomes of unicellular (Fleischmann *et al.*, 1995) and, later, multicellular (C. elegans Sequencing Consortium, 1998) organisms had been sequenced. The major breakthrough in this area took place in 2001, when the first draft of the human genome was released (Venter *et al.*, 2001). Two years later, 100 complete genome sequences had been submitted to databases (Janssen *et al.*, 2003), and the recent development in sequencing technology makes it difficult to estimate the current number of completed genomes.

5.2.2.1 Microarrays

Genome-wide studies are conducted with microarrays (Schena *et al.*, 1995). Usually, these arrays are used to probe the differences in gene expression after exposure to some substance, bacterium, or condition. They consist of either DNA fragments from the investigated organism or oligonucleotides, which are short sequences of synthetic nucleotides (DNA or RNA) corresponding to the gene sequence of the investigated organism. In both cases, the idea is to capture labeled complementary DNA molecules, which are synthesized from the mRNAs of individual genes, to a single spot on the array. When complementary DNA molecules from different conditions are labeled with different colors, the proportion of color on a spot reflects the change in the expression of that gene between the studied conditions. A single microarray may contain tens of thousands of spots.

Microarrays can be used at several stages in the drug development process (Gerhold *et al.*, 2002). First, microchips are used at the drug discovery stage in order to find genes or pathways that are suitable drug targets. Usually, cell lines are infected or altered to resemble a disease, and changes in the gene expression profile are monitored. The goal is to find genes or gene pathways that are crucial for the disease process, to use these genes or their products as targets, and to screen or design molecules against them in order to stop the progression of the disease. The target validation stage confirms that the selected target is, in fact, crucial for the disease, and microchips are used to monitor the changes that

take place when the gene that encodes this target is “turned off”. This data is used to find the targets that affect only the same genes or gene pathways as the disease and, because of this, hopefully do not cause any side effects (Gerhold *et al.*, 2002). In the lead molecule optimization process, microarrays are used in the characterization, optimization, and finally, selection of the best lead molecule by studying the expression profiles caused by the lead candidate compounds. This selection process also involves studies on metabolism and toxicology. In these studies, the expression profiles of potential drug compounds are compared to the profiles of well-known compounds with desirable/undesirable metabolism and toxicology profiles.

The utilization of microarrays in drug development is not limited to the pre-clinical phase. Disease versus control studies can also be utilized in the clinical research (Weeraratna *et al.*, 2004), and there are clinical screening programs where microarrays are used to predict the best possible treatment (Schubert, 2003). In addition, microarrays are in themselves valuable diagnostic, prognostic, and predictive tools (Pusztai *et al.*, 2003), but their full potential in clinical trials can only be utilized after the problems related to biological, methodological, and statistical issues have been resolved (Olson, 2004).

5.2.3 Proteomics

The word “proteome” refers to all the proteins produced by an organism, and “proteomics”, first described in 1969 (Macko and Stegemann, 1969), is the study of their expression, structure, and function (Persidis, 1998). Proteomics is the next logical step after genomics, since most of the cell functions take place at the protein level. The changes at the protein level are more relevant for the cell, but there are no methods able to quantify and identify proteins on the global scale, although there has been promising development in the proteomics methodology in the recent years (Arthur, 2003). The active interest in proteomics is due to the fact that changes at the transcription level do not necessarily show up at the protein level. There are several reasons for this, including the rapid degradation or inefficient translation of messenger RNA, and the major advance of proteomics, compared to microarrays, is that it can also detect protein isoforms resulting from post-translational modifications, proteolytic processing, or alternative mRNA splicing caused by the given condition.

5.2.3.1 Two-dimensional polyacrylamide gel electrophoresis

The basic idea of proteomics is first to separate and then to identify individual proteins. Today, the most widely used proteomics tool is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) coupled with mass spectrophotometry (MS). In this system, proteins are separated as spots in a gel using 2D-PAGE and identified with MS. Separation on the first dimension (horizontal) is done using a tube gel or an immobilized pH gradient strip, and it is based on the isoelectric point of the protein (Klose and Kobalz, 1995). Separation on the second dimension (vertical) is based on the mass of the protein, and it is done using electrophoresis. After these separations, the proteins are visualized with a stain. Individual protein spots are then cut off from the gel and identified in MS. There are problems related to solubility and the separation of structurally similar proteins, but this method can give valuable information about the relevant changes inside the cell.

Pharmaceutical industry has taken a cautious approach towards proteomics, mainly because of the recent large investments in the genomic and HTS area (Hanash, 2003). Currently, proteomics is used mostly in expression proteomics studies in drug / biomarker discovery as well as in toxicoproteomics. Expression proteomics tries to identify and characterize proteins that are expressed differently in cell lines, tissues, or organisms under given condition. The aim of these studies is to find proteins that can be used as drug targets (Van Eyk, 2001) or diagnostic biomarkers (Petricoin *et al.*, 2002). In toxicity testing, proteomics is used to identify altered proteins or protein pathways after drug treatment (Wetmore and Merrick, 2004).

6 AIMS OF THE STUDY

The overall aim of this thesis was to find compounds and specific target molecules that could be used in the treatment of *C. pneumoniae* infections. The screening of anti-chlamydial compounds was done using focused libraries t suitable for an academic environment. Specific target molecules were searched using genomics and proteomics.

The specific aims of the study were to:

1. Search for potent compounds against *C. pneumoniae* infection *in vitro*

2. Develop a new high-throughput screening method and screen natural and synthetic compounds for molecules that inhibit *C. pneumoniae* growth.

3. Study the effects of *C. pneumoniae* infection on the host cell's gene and protein expression at different stages of the infection.

4. Identify new potential drug targets for the treatment of *C. pneumoniae* infection using proteomics and genomics data.

5. Validate the significance of the discovered potential drug target genes by silencing them with specific siRNA molecules and measuring the effect of silencing on the number of *C. pneumoniae* particles.

7 EXPERIMENTAL

7.1 Materials

7.1.1 Cell line and chlamydia strains (**I -V**)

The human epithelial cell line (HL) (Cles and Stamm, 1990) used in this study was cultivated in RPMI 1640 medium (BioWhittaker Europe, Verviers, Belgium) supplemented with 7.5% of heat-inactivated fetal bovine serum of South American origin (BioWhittaker Europe), 2mM L-glutamine, and 20 μg/ml streptomycin. The cells were kept at 37°C and 95% humidity in an atmosphere containing 5% CO2.

The chlamydial strains used in this study (**Table 1**) were purified from infected cells by Urografin (Schering Ag, Berlin, Germany) gradient ultracentrifugation. Chlamydial EBs were stored in sucrose–phosphate–glutamic acid buffer (0.2M sucrose, 3.8mM KH_2PO_4, 6.7mM Na_2HPO_4, and 5mM L-glutamic acid, pH 7.4) at -70°C.

Table 1. Chlamydial strains

Bacterial strains	**Description**	**Article**	**Source**
C. pneumoniae K7	*C. pneumoniae* isolate Kajaani 7	I, II, III	(Ekman *et al.*, 1993a)
C. pneumoniae CWL-027	*C. pneumoniae* isolate	IV, V	ATCC number VR-1310

7.1.2 Pure compounds and synthetic derivatives (**I**)

The plant-derived pure compounds and synthetic derivatives used in this study are presented in **Table 2**. All compounds were dissolved in dimethyl sulfoxide (DMSO); the final concentration of DMSO during the experiments was 0.5%, and it did not affect the host cells or the infection.

Table 2. Plant-derived pure compounds and synthetic derivatives.

Compound	Description	Article	Source
Acacetin	Flavones	I	Carl Roth GmbH
Apigenin	Flavones	I	Fluka
Daidzein	Isoflavones	I	Extrasynthese
Daidzin	Isoflavones	I	Extrasynthese
Flavone	Flavones	I	Carl Roth GmbH
Genistein	Isoflavones	I	Extrasynthese
Genistin	Isoflavones	I	Extrasynthese
Isorhamnetin	Flavonols	I	Extrasynthese
Kaempherol	Flavonols	I	Extrasynthese
Luteolin	Flavones	I	Extrasynthese
Luteolin- 7 –glucoside	Flavones	I	Extrasynthese
Luteolin-3',7-glucoside	Flavones	I	Extrasynthese
Morin	Flavonols	I	Carl Roth GmbH
Myricetin	Flavonols	I	Extrasynthese
Naringenin	Flavanones	I	Sigma
Naringin	Flavanones	I	Extrasynthese
Procyanidin B l	Catechins	I	Extrasynthese
Procyanidin B2	Catechins	I	Extrasynthese
Quercetin	Flavonols	I	Merck
Quercitrin	Flavonols	I	Carl Roth GmbH
Rhamnetin	Flavonols	I	Extrasynthese
Rutin	Flavonols	I	Merck
Vitexin	Flavones	I	Extrasynthese
Vitexin-2"-O-rhamnoside	Flavones	I	Extrasynthese
Benzoic acid	Phenolic acids	I	Merck
Caffeic acid	Phenolic acids	I	Sigma
Dodecyl gallate	Gallates	I	Fluka
Ferulic acid	Phenolic acids	I	Extrasynthese
Gallic acid	Phenolic acids	I	Sigma
Methyl gallate	Gallates	I	Fluka
Octyl gallate	Gallates	I	Extrasynthese
Propyl gallate	Gallates	I	Sigma
Syringic acid	Phenolic acids	I	Sigma
Resveratrole	Stilbene	I	Extrasynthese
(+)-Catechin	Catechins	I	Sigma
Coumarin	Natural coumarins	I	Sigma
(-)-Epicatechin	Catechins	I	Sigma
(-)-Epicatechin gallate	Gallates	I	Extrasynthese
(-)-Epigallocatechin	Catechins	I	Extrasynthese
Methoxy psoralen	Natural coumarins	I	Fluka
Scopoletin	Natural coumarins	I	Sigma
Umbelliferone	Natural coumarins	I	Sigma
Xanthotoxin	Natural coumarins	I	Extrasynthese
2 '-methoxy-a-naphtoflavone	Synthetic flavonoids	**I**	ICC chemical corporation
3-(2-benzoxazoyl)umbelliferone	Synthetic coumarins	I	Fluka
3 -(a-acetonylbenzyl)-4-hydroxycoumar	Synthetic coumarins	I	Sigma

3-benzoylbenzo(F)coumarin	Synthetic coumarins	I	Acros
4-methyl-3-phenylcoumarin	Synthetic coumarins	I	Sigma
6-methylcoumarin	Synthetic coumarins	I	Extrasynthese
6,2 , -dimethoxyflavone	Synthetic flavonoids	I	Sigma
6,8-dibromocoumarin-3-carboxylic acid	Synthetic coumarins	I	Avocado
7-diethylamino-3-thenoylcoumarin	Synthetic coumarins	I	Acros
Alpha-naphtoflavone	Synthetic flavonoids	I	Acros
Coumarin 30	Synthetic coumarins	I	MP Biomedicals
Coumarin 102	Synthetic coumarins	I	Acros
Coumarin 106	Synthetic coumarins	I	Acros
Rotenone	Synthetic flavonoids	I	Acros

7.1.3 Antimicrobial agents (**II**)

To determine the minimal inhibitory concentration (MIC) of *C. pneumoniae*, fresh solutions of antimicrobial agents (**Table 3**) were prepared for each experiment. All agents were dissolved in DMSO; the final concentration of DMSO during the experiments was 0.2%, and it did not affect the host cells or the infection.

Table 3. Antimicrobial agents used in this study.

Antimicrobial agents	**Description**	**Article**	**Source**
Erythromycin	Commercial antibiotic	II	Sigma
Streptomycin	Commercial antibiotic	II	Sigma
Rifampicin	Commercial antibiotic	II	Sigma
Doxycycline	Commercial antibiotic	II	ICN Biomedicals
Ofloxacin	Commercial antibiotic	II	ICN Biomedicals
Minocycline	Commercial antibiotic	II	ICN Biomedicals
Ciprofloxacin	Commercial antibiotic	II	ICN Biomedicals

7.1.4 Synthetic compounds (**III**)

The X-ray crystal structure of a potential *C. pneumoniae* target protein homologue (*Bacillus subtilis* RNA methyltransferase; coded by the *ermC* gene) was used in the virtual screening of approx. 300 000 commercial compounds in order to create a targeted library against the selected *C. pneumoniae* target protein (dimethyladenosine transferase; coded by the *ksgA* gene). After the virtual screening process, the 2000 best binding molecules, as ranked by the FlexX program, were analyzed visually, and the 33 most promising molecules (**Table 4**.) with good docking orientations were ordered from Specs (Delft, Netherlands) and Maybridge (Cornwall, England).

Table 4. The compounds S1 – S12 were ordered from Specs and MB1 – MB21 from Maybridge.

Abbreviation	International Union of Pure and Applied Chemistry (IUPAC) names
S1	2-{4-amino-6-[(4-chlorophenyl)amino]-1,3,5-triazin-2-yl}-4-chlorophenol
	N-1,3-benzodioxol-5-yl-2-[(5-methyl-1*H*-benzimidazol-2-yl)thio]acetamide
S2	2-(1*H*-benzimidazol-2-ylthio)-*N*-(3-hydroxyphenyl)acetamide
S3	*N*-(3-fluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide
S4	2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)-*N*-phenylacetamide
S5	*N*-1,3-benzodioxol-5-yl-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide
S6	*N*-1,3-benzodioxol-5-yl-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide
S7	2-[(1*H*-benzimidazol-2-ylmethyl)thio]-3*H*-imidazo[4,5-*b*]pyridine
S8	*N*-[3-(acetylamino)phenyl]-4-methoxybenzamide
S9	3-[(2,5-dichlorophenoxy)methyl]-*N*-2,3-dihydro-1,4-benzodioxin-6-ylbenzamide
S10	*N*-(2,4-difluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide
S11	*N*-(4-fluorophenyl)-2-(3*H*-imidazo[4,5-*b*]pyridin-2-ylthio)acetamide
S12	*N*-1,3-benzodioxol-5-yl-2-[(5-methoxy-1*H*-benzimidazol-2-yl)thio]acetamide
MB1	*N*-(1,3-benzodioxol-5-ylmethyl)-6-phenylthieno[3,2-*d*]pyrimidin-4-amine
MB2	*N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-fluorobenzamide
MB3	*N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-methylbenzamide
MB4	*N*-[3-(1*H*-benzimidazol-2-yl)phenyl]thiophene-2-carboxamide
MB5	*N*-[3-(1*H*-benzimidazol-2-yl)phenyl]benzamide
MB6	*N*-[3-(1*H*-benzimidazol-2-yl)phenyl]-4-methoxybenzamide
MB7	*N*-2,3-dihydro-1,4-benzodioxin-6-ylbenzamide
MB8	*N*-2,3-dihydro-1,4-benzodioxin-6-yl-4-methoxybenzamide
MB9	*N*-(4-morpholine-4-aminobenzyl)-3,4-dihydro-2*H*-1,5-benzodioxepine-7-carboxamide
MB10	9-{2-[(1,3-benzodioxol-5-ylmethyl)amino]ethyl}-9*H*-purin-6-amine
MB11	*N*-{4-[(pyrimidin-2-ylamino)sulfonyl]phenyl}acetamide
MB12	*N*'-[(anilinocarbonyl)oxy]-2,1,3-benzoxadiazole-5-carboximidamide
MB13	*N*'-({[(3-chlorophenyl)amino]carbonyl}oxy)-2,1,3-benzoxadiazole-5-carboximidamide
MB14	2-[(2,1,3-benzoxadiazol-5-yloxy)methyl]-*N*-(4-chlorophenyl)-1,3-thiazole-4-carboxamide
MB15	2-[(2,1,3-benzoxadiazol-5-yloxy)methyl]-*N*-(4-nitrophenyl)-1,3-thiazole-4-carboxamide
MB16	N^2,N^4-diphenyl-1,3,5-triazine-2,4,6-triamine
MB17	*N*-benzyl-9*H*-purin-6-amine
MB18	*N*-(2-furylmethyl)-9*H*-purin-6-amine
MB19	*N*-(3-nitrophenyl)-1,3,5-triazine-2,4-diamine
MB20	4-chloro-*N*-(2-{[(6-nitro-4*H*-1,3-benzodioxin-8 yl)methyl]thio}phenyl) benzamide
MB21	*N*-(2-{[(6-nitro-4*H*-1,3-benzodioxin-8-yl)methyl]thio}phenyl)thiophene-2-carboxamide

7.1.5 Microarrays and scanner (**IV**)

RNA extracted from infected and uninfected cells was labeled with the RNA labeling kit (Agilent Fluorescent Direct Labeling Kit, Agilent, USA) and hybridized to the Agilent Technologies (Santa Clara, CA 95051, United States) human 1A oligo microarrays (V2). The microarrays were scanned with a GenePix 4200 autoloader scanner (Axon instruments).

7.2 Methods

7.2.1 Mycoplasma detection (**I, IV, V**)

Mycoplasma detection from the HL cell line was done with either the mycoplasma Plus PCR Primer Set (Stratagene, California, USA) (**IV**) or Mycoplasma PCR ELISA (Roche Diagnostics, Mannheim, Germany) (**I** and **V**)

7.2.2 Infection protocol (**I - V**)

In the 6-well plate experiments 3.5 x 10^6 (IV) or 3 x 10^6 HL cells/well (**V**) were allowed to attach to the bottom of the wells and the cell monolayers were infected 24 h later with of 7 x 10^6 (**IV**) or 6 x 10^6 (**V**) IFU/well (CWL-029 strain), using 550 x g centrifugation for 60 min at room temperature. After infection, the cells were washed once with phosphate-buffered saline, fresh culture medium was added to the wells, and the plates were placed into a 5% CO_2 atmosphere at 37°C for different periods of time. This protocol was used when infected cell were utilized in genomics and proteomics experiments.

In the 24-well plate experiments 4.5 x 10^5 cells/well were allowed to attach to the coverslips at the bottom of the wells. Confluent cell monolayers were infected 24 h later with an infection medium supplemented with 0.5 μg/ml of cycloheximide, an 80-s ribosome inhibitor, and 10 x 10^3 (**I**) or 8 x 10^4 (**II**) IFU/well (K7 strain) using 550 x g centrifugation for 60 min at room temperature. After infection, the cells were washed once with phosphate-buffered saline, fresh culture medium containing cycloheximide and the tested compound was added to the wells, and the plates were placed into a 5% CO_2 atmosphere at 37°C for 72 h. This protocol was used when *C. pneumoniae* inclusions were counted after IF staining.

In the 96-well plate experiments, 6 x 10^4 (**II** and **IV**) or 5 x 10^4 (**III**) cells/well were allowed to attach to the clear-bottomed 96-well plates. Confluent cell monolayers were infected 24 h later with an infection medium supplemented with 0.5 μg/ml of cycloheximide **(II** and **III**), an 80-s ribosome inhibitor, and 1.4 x 10^4 (**II**), 8 x 10^4 (**III**) or 1.2 x 10^5 (**IV**) IFU/well (K7 in **II** and **III** and CWL-029 in **IV**) using 550 x g centrifugation for 60 min at room temperature. After infection, the cells were washed once with phosphate-buffered saline, fresh culture medium containing cycloheximide and the tested compound **(II** and **III**) or siRNA molecule (**IV**) was added to the wells, and the plates were placed into a 5% CO_2 atmosphere at 37°C for 72 h. This protocol was used in the TR-FIA measurements.

7.2.3 Cytotoxicity testing (**I, III, IV**)

The Lactate dehydrogenase (LDH) kit (CytoTox, Promega, USA) (**I** and **III**) and the Adenosine triphosphate (ATP) kit (ATPlite 1step, PerkinElmer Life and Analytical Sciences, USA) (**IV**) were used according to the manufacturers' instruction in order to detect the possible cytotoxicity of the tested compounds (**I** and **III**) or siRNA molecules (**IV**).

7.2.4 Nucleic acid isolation and q-PCR (**II, IV**)

DNA was extracted from infected and uninfected cells using a DNA mini kit (Qiagen), according to the manufacturer's instructions (**II**).

Q-PCR with SYBR Green dye was used for the relative quantification of *C. pneumoniae* DNA in the samples, using a LightCycler instrument (Roche Biochemicals) and the FastStart DNA Master SYBR Green I kit (Roche Biochemicals), VD-4 primers, and the cycling conditions described by Tondella *et al*. (Tondella *et al.*, 2002) (**II**).

Total RNA was extracted from infected and uninfected cells with the RNeasy kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions (**IV**).

The expression of the selected genes was confirmed with q-PCR. 2 µg of RNA was reverse-transcribed using the TagMan Reverse Transcription Reagents, Abgene - Absolute SYBR Green Rox Mix and ABI Prism 7000 Sequence Detection System for q-PCR (**IV**). The primers that were used in q-PCR are shown in **Table 1 in IV**.

7.2.5 *C. pneumoniae* detection (**I – IV**)

The immunofluorescence staining of C. *pneumoniae* inclusions was done 72 h p.i. The culture medium was removed from the wells, and the coverslips were washed once with PBS and fixed in methanol for 10 min. The coverslips were allowed to dry, and the chlamydial inclusions were stained directly with fluorescein isothiocyanate (FITC)-conjugated chlamydial genus-specific antibody (PathFinder1 Chlamydia Culture Confirmation System, Bio-Rad S.A., WA, USA). The stained inclusions were

examined under a fluorescence microscope (Nikon ECLIPSE TE300 inverted microscope with TE-FM epi-fluorescence attachment, Tokyo, Japan) (**I** and **II**).

In the TR-FIA protocol, infected HL cells were fixed in 96-well plates and labeled with a genus-specific murine monoclonal antibody (Argene SA, Varilhes, France) containing an europium label (PerkinElmer/Wallac Oy, Turku, Finland). After 30-min labeling at 37 °C, the plates were washed six times with a Biohit BW50 platewasher (Biohit Plc, Finland) using 300 μL of Wallac DELFIA Wash Solution (Perkin-Elmer Life and Analytical Sciences/Wallac Oy) each time. An aliquot of 100 μL of DELFIA Enhancement Solution (PerkinElmer) was added to each well, and the plates were shaken for 5 min at low speed on a DELFIA Plateshake (PerkinElmer). Signals associated with the *Chlamydia* antibody were measured from the wells with a Wallac Victor2 multilabel counter (PerkinElmer) (**II**, **III** and **IV**).

7.2.6 Virtual screening (**III**)

The peptide chain from the X-ray crystal structure 1QAO was used as the target protein. The structure was pre-processed in the Sybyl software, version 6.5 (Tripos Inc., St. Louis, USA). The atom types of carboxylic acid oxygens were set to O.co2 and those of basic nitrogens to N.4. Hydrogens were added to the model, and their orientation was optimized using Tripos force field energy minimization, while all non-hydrogen atoms were not allowed to move (**III**).

The ligand position in 1QAO was used to define the active site cavity. Default FlexX (Sybyl 6.5) parameters were used. The 2000 top ranking molecules were further inspected visually to exclude any structures with improbable docking orientation, and 33 molecules with good docking orientations were chosen for biological testing (**III**).

7.2.7 Gene silencing **(IV**)

SiRNAs molecules were ordered from Qiagen (Valencia, CA, USA) (**IV**). Gene silencing was executed using Qiagens HiPerFect Transfection Reagent in line with the HiPerFect Transfection Reagent Handbook instructions. Two siRNAs were tested for each gene. The results were validated in the ABI Prism 7000 Sequence Detection System (Applied Biosystem, USA) using “Hs_GAPD_5 HP Validated siRNA” as a positive control and “control (non-silence) siRNA” as a negative control.

8 RESULTS AND DISCUSSION

The work for this thesis is part of a larger ongoing research project (**Fig. 3**), which aims to develop new specific drug molecules and diagnostic tests for *C. pneumoniae* infection.

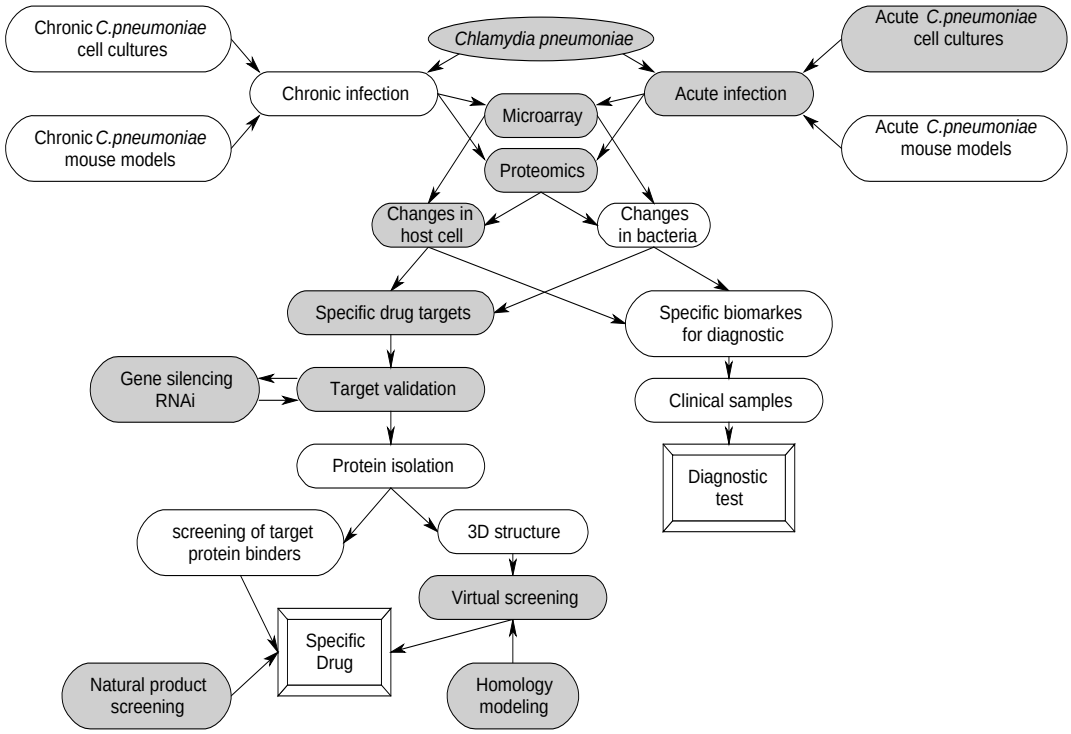

Figure 3. A schematic presentation of the entire research project. The grey boxes are included in this thesis.

8.1 Screening of active compounds (**I, II, III**)

The search for anti-chlamydial compounds was started by screening a group of natural and naturally derived pure compounds, which were already known to have antimicrobial properties (**I**). This screening was done by using a traditional staining and inclusion counting method, which is laborious and time-consuming. These reasons led to the development of a new screening assay, which is more automated and suitable for high-throughput screening (**II**). This new assay was used to screen a small targeted

library, built with the help of a structural homolog of the selected target, in order to find anti-chlamydial activity among the synthetic compounds (**III**).

8.1.1 Natural products (**I**)

A group of natural and natural-based products (57 in total) were screened against *C. pneumoniae*, and the anti-chlamydial activity of these compounds was determined by counting inclusions after immunofluorescence staining (**Table 5**). The *in vitro* minimum inhibitory concentration (MIC) was also determined for some of the most active compounds (**Table 1 in I**).

Highly active compounds were found from several groups, but gallates were clearly the most active. The antibacterial activity of gallates has already been demonstrated in other species, and the length of the alkyl group has been found to increase this activity (Kubo *et al.*, 2003). This was also the case in our studies (**Table 1 in I**), since the MIC value for dodecyl gallate was 18 μM (the length of the alkyl group was 12 carbons) and that for octyl and methyl gallate 29 μM (the lengths of the alkyl groups were 3 and 1 carbon, respectively). Inactive compounds were also seen in several groups, but most compounds with 0% inhibition occurred among synthetic coumarins (**Table 5**). Natural coumarins, on the other hand, were quite active against *C. pneumoniae*, which is not surprising, since they are already known to have antibacterial activity (Ojala *et al.*, 2000). The overall high antichlamydial activity of the phenolic compounds included in our study is in good agreement with corresponding studies using different phenolics and bacteria (Puupponen-Pimia *et al.*, 2005) (Rauha *et al.*, 2000).

The antimicrobial properties of natural compounds often originate from their ability to interact with biological membranes (Maillard, 2002). In order to see whether the compounds in our study have the capacity to retain and/or penetrate membranes, HL cells were pre-incubated for 24 h with the investigated compound before the actual *C. pneumoniae* infection.

Table 5. The average inhibition percentages of plant-derived phenolic compounds and some synthetic derivatives against *C. pneumoniae* (50 µM concentration; n=4 or more). The following categories of activity were used: Highly active = 85-100% inhibition, compared to DMSO controls; Active = 50-84%; Moderately active = 30-49%; Inactive = < 30%. Abbreviations of the compound groups; Catechins (A), Gallates (B), Flavanones (C), Flavones (D), Flavonols (E), Isoflavones (F), Natural coumarin (G), Phenolic acids (H), Stilbene (I), Synthetic coumarins (J) and Synthetic flavonoids (K).

Highly active

Compound	Inhibition %
Dodecyl gallate (B)	100
Methyl gallate (B)	100
Octyl gallate (B)	100
Propyl gallate (B)	100
(-)-Epicatechin gallate (B)	85
Acacetin (D)	100
Apigenin (D)	100
Luteolin (D)	100
Flavone (D)	90
Morin (E)	100
Myricetin (E)	100
Rhamnetin (E)	100
Quercetin (E)	90
Methoxy psoralen (G)	100
Scopoletin (G)	96
Xanthotoxin (G)	94
7-diethylamino-3-thenoylcoumarin (J)	100
Coumarin 106 (J)	100
2 '-methoxy-a-naphtoflavone (K)	100
Rotenone (K)	100
Alpha-naphtoflavone (K)	92

Active

Compound	Inhibition %
(+)-Catechin (A)	76
(-)-Epicatechin (A)	75
(-)-Epigallocatechin (A)	58
Naringin (C)	66
Isorhamnetin (E)	70
Quercitrin (E)	50
Genistein (F)	60
Daidzein (F)	51
Umbelliferone (G)	75
Caffeic acid (H)	78
Resveratrole (I)	54
6-methylcoumarin (J)	71
Coumarin 102 (J)	63
3-benzoylbenzo(F)coumarin (J)	62
Coumarin 30 (J)	50
6,2 , -dimethoxyflavone (K)	73

Moderately active

Compound	Inhibition %
Procyanidin B l (A)	30
Luteolin-3',7-glucoside (D)	45
Rutin (E)	46
Genistin (F)	37
Benzoic acid (H)	44
Syringic acid (H)	32

Inactive

Compound	Inhibition %
Procyanidin B2 (A)	0
Naringenin (C)	16
Luteolin- 7 –glucoside (D)	23
Vitexin-2"-O-rhamnoside (D)	11
Vitexin (D)	3
Kaempherol (E)	15
Daidzin (F)	0
Coumarin (G)	28
Ferulic acid (H)	14
Gallic acid (H)	27
3-(2-benzoxazoyl)umbelliferone (J)	28
3 -(a-acetonylbenzyl)-4-hydroxycoumarin (J)	0
4-methyl-3-phenylcoumarin (J)	0
6,8-dibromocoumarin-3-carboxylic acid (J)	0

The infection was done according to the infection protocol, and the infected cells were grown in the infection medium without the test compound for 72 h (**Fig. 4**). These results confirmed the previously discovered characteristic of these compounds (Tammela *et al.*, 2004), demonstrating that some of them have the ability to accumulate inside cells or into cell membranes and reduce the number of *C. pneumoniae* inclusions, even when present only prior to infection.

To analyze if *C. pneumoniae* could be totally eliminated from the cell cultures, the infectivity of the possible *C. pneumoniae* progenies after exposure to four active compounds was determined (**Fig. 4**). Infected HL cells cultivated with the tested compound were mechanically disrupted after 72 h incubation, repassaged on freshly prepared confluent HL cell layers, and incubated for 72 h without the compounds.

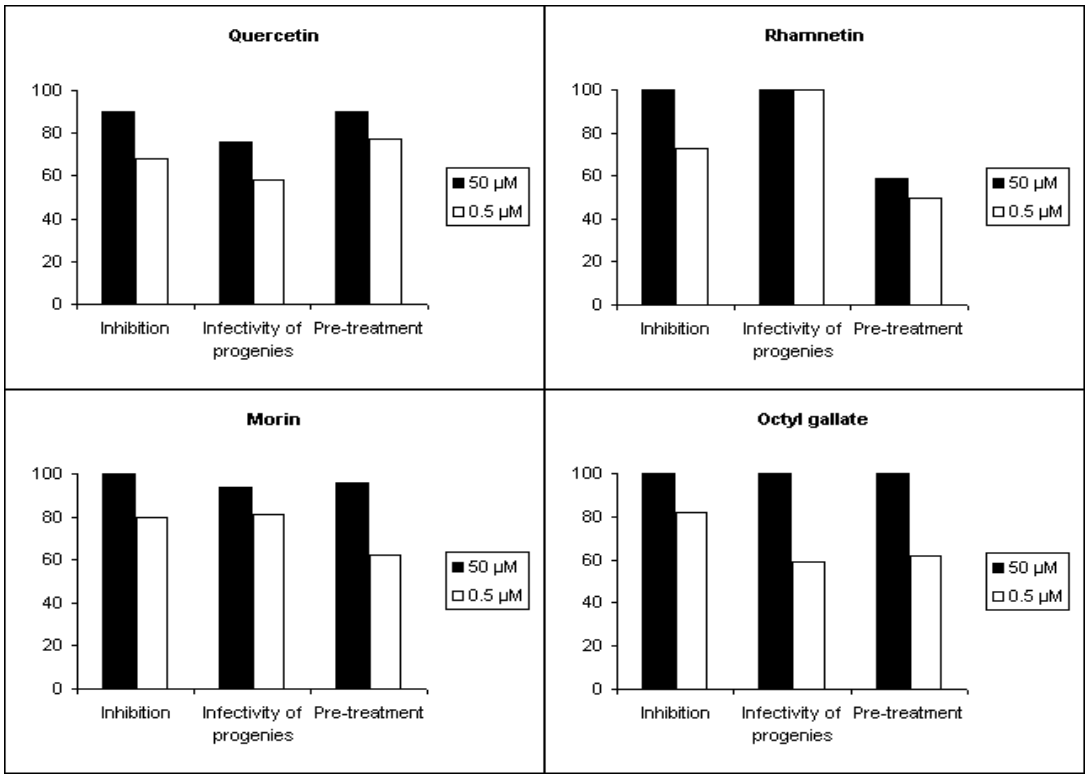

Figure 4. Effect of pre-treatment and continuous treatment on the formation of chlamydial inclusions and the yield of infectious chlamydial particles. Values are average inhibition percentages, compared to untreated controls 72 h p.i.

The results showed that quercetin, morin, and octyl gallate do not affect the infectivity of newly formed EBs, since inclusions could be found after the repassage. Rhamnetin, on the other hand, seemed to have an effect on newly formed EBs or it may block the conversion of RBs back to EBs at the end of the developmental cycle, since there was 100% inhibition after the repassage, although there were inclusions left after the initial 0.5 mM treatment.

The antichlamydial activity of some of the most active compounds of this study, i.e. quercetin, luteolin, and octyl gallate, has also been evaluated *in vivo* (Törmakangas *et al.*, 2005). In these experiments, mice were treated for three days before and for ten days after *C. pneumoniae* inoculation in order to see how these compounds modify the course of the infection. Lung tissue analysis, assessment of the inflammatory response, and *C. pneumoniae* antibody levels demonstrated a clear beneficial effect for luteolin, no significant effect for octyl gallate, and an unbeneficial effect for quercetin, which was speculated to be caused by incorrect dosage and an improper route of administration. The beneficial effect of luteolin was due to the decreased number of *C. pneumoniae* particles in the lung tissue and the prevention of lung inflammation. Although this study validated our observation about the antichlamydial effect of luteolin, it simultaneously demonstrated the poor correlation between the *in vitro* and *in vivo* experiments for quercetin and octyl gallate.

Structural comparisons between active and inactive compounds enabled detection of some structural characteristics that differentiated these compounds in some compound groups. Flavones and flavonols have the same basic structure, and both groups contain active and inactive compounds. In these compound groups, structure is related to activity in such a way that all the compounds with 50% or less inhibition contain a sugar moiety or moieties as substituents, whereas none of the more active compounds (over 70% inhibition) contain sugars.

8.1.2 Development of a new screening assay for *C. pneumoniae* (**II**)

Conventional immunofluorescence staining (IF staining) is a widely used method in *C. pneumoniae* research, but it is unsuitable for the screening of even medium-sized compound libraries. We set out to develop a new cell-based assay that could be automated and used in the 96-well format. This new assay, named time-resolved fluorometric immunoassay (TR-FIA), is based on time-resolved fluorometry. TR-FIA utilizes eurobium-conjugated antibody, which enables reliable separation between labeled *C.*

pneumoniae particles and the background, since eurobium label has a considerably longer half-life than normal (background) fluorescence (Soini and Kojola, 1983).

The optimal conditions for TR-FIA were defined in accordance with Zhang *et al* (Zhang *et al.*, 1999), using statistical parameters such as the signal-to-background (S/B) ratio, the signal-to-noise (S/N) ratio, and the screening window coefficient calculated from control sample data (Z' factor). The S/B ratio measures the separation between the signal and the background, whereas the S/N ratio also takes into account information regarding data variation. The Z' factor measures the ratio between the signal window and the dynamic range of the signal and is widely used to assess the quality of an assay for high-throughput screening. These three parameters were used to optimize the most critical factors of the new assay, namely antibody concentration, incubation time, and *C. pneumoniae* concentration. Comparison of the TR-FIA results to the IF staining and q-PCR was used to define the reliability of the assay.

After optimization, the repeatability of TR-FIA was very good compared to other cell-based assays. Plate-to-plate variation was only 8.9% (n=3) and day-to-day variation 9.0% (n=3). In addition, when optimized conditions were used, the statistical parameters S/B, S/N, and Z' were satisfactory for a cell-based assay (6.46, 6.85, and 0.39, respectively).

In general, the results obtained with TR-FIA were similar to the results of IF staining and q-PCR assay (**Table 3 in II**), but q-PCR occasionally gave slightly lower MIC values than TR-FIA or IF staining. This can be explained by the more sensitive nature of the PCR technology, which enables it to react faster to the decreasing number of *C. pneumoniae* particles. When the size of inclusions is decreased prior to total inhibition, as in the case of ciprofloxacin (Hammerschlag *et al.*, 1992a), TR-FIA is also able to detect this decrease. Besides these small differences, the only significant exception was seen for penicillin, which gave a clearly lower MIC value with q-PCR than TR-FIA or IF staining. This difference can be explained by penicillin's mechanism of action, which is unique among the antimicrobial agents used in this study. Penicillin blocks peptiglycan synthesis, which in *C. pneumoniae* is related to chlamydial cell division (McCoy *et al.*, 2003) and possibly the conversion of EBs from RBs at the end of the development cycle (Chopra *et al.*, 1998). This being the case, penicillin reduces the number of chlamydial particles inside the inclusions, which would be seen as inhibition in q-PCR, but does not reduce the number of inclusions, which means that no reduction can be seen in IF staining. Furthermore, the blockage of RB binary fission results in abnormally large RBs (Matsumoto and

Manire, 1970), which contain more lipopolysaccharide, the antigen of IF staining, and TR-FIA antibody. This explains why TR-FIA does not detect the reduction of the chlamydia particles.
In conclusion, it can be said that TR-FIA gives reliable and reproducible results and is an excellent tool for the purpose it was designed for, i.e. primary screening of new antichlamydial compounds.

8.1.3 Synthetic compounds (**III**)

In addition to natural compounds, we also wanted to search for antichlamydial agents among synthetic compounds. Overall, synthetic libraries are large, and synthetic compounds have less biological activity than natural or natural-based compounds, but on the good side, individual synthetic compounds are cheaper and much easier to produce.

Since we lacked the resources to screen large compound libraries, we decided to create a targeted library by virtually screening molecules that show affinity to the 3D structure of a selected *C. pneumoniae* target protein. The only problem in this case was that there are no publicly available 3D structures for any *C. pneumoniae* proteins. We tried to solve this problem by utilizing the inability of virtual screenings to reliably discriminate ligand binding between closely related protein isoforms (Schapira *et al.*, 2003). We wanted to find the 3D structure of a close homolog of the *C. pneumoniae* target protein and to use it to screen molecules that would be more likely to bind to the actual *C. pneumoniae* target. In this way we could create a targeted library and reduce the number of molecules in the *C. pneumoniae* assay, without actually doing the time-consuming and expensive crystallization of the target protein.

We started to search databases in order to find the *C. pneumoniae* target protein and a closely related non-human protein with a crystallized 3D structure and a highly similar active site area. We did not create a comparative model for the *C. pneumoniae* target protein, since we wanted to see if a high level of similarity in the amino acid sequence, and especially the binding pocket, is enough to guide the creation of a small library of active compounds. Extensive data mining produced a promising target-homolog pair. The target was *C. pneumoniae* dimethyladenosine transferase (coded by the *ksgA* gene) and the crystallized homolog *Staphylococcus aureus* rna methyltransferase (coded by the *ermC* gene), which causes antibiotic resistance in many bacteria (but not in *C. pneumoniae*). These methyltransferases have highly similar active sites (**Table 2 in III**), and they bind their ligands in a similar way (Seppälä *et al.*, 1998). In addition, in the International Union of Biochemistry and Molecular Biology (NC-IUBMB) classification, they belong to the same subclass: rRNA (adenine-N^6-)-

methyltransferase (EC: 2.1.1.48), whose indispensable functions are related to ribosomal structure and ribosomal methylation.

The 3D structure of *S. aureus* RNA methyltransferase was used to virtually screen for a compound from two large Internet databases. These commercial databases, Specs (Delft, Netherlands: http://www.specs.net/) and Maybridge (Cornwall, England: http://www.maybridge.com/), contained over 300 000 molecules, and they were screened *in silico* in order to create a small targeted library, which should be highly active against *C. pneumoniae*. After the screening process, the 2000 best binding molecules, as ranked by the FlexX program, were analyzed visually, and 33 molecules were purchased and their anti-chlamydial properties were tested using TR-FIA (**Fig. 5**).

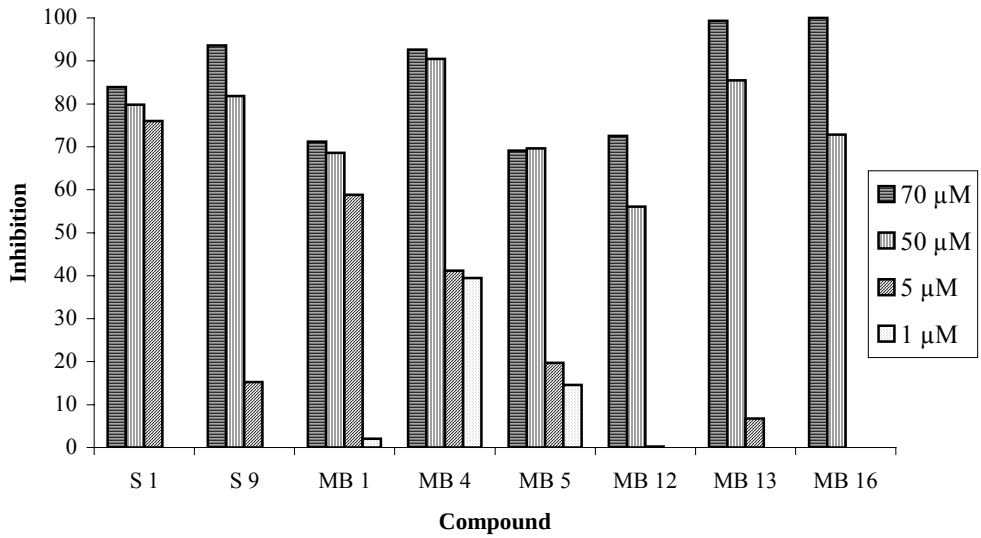

Figure 5. The most effective anti-chlamydial synthetic compounds.

Eight out of the 33 tested compounds produced over 50% inhibition at 50 µM concentration. This activity percentage (24.2%) is extremely high compared to random libraries in general (Zolli-Juran *et al.*, 2003), (Lewis *et al.*, 2004), (Carey *et al.*, 2004), and (Pilger *et al.*, 2004). In the screening of a natural and natural-based compound library (**I**), 35 out of 57 compounds (61.4%) produced over 50% inhibition at the same concentration. Hence, although the virtual screening with the 3D structure of a close homolog increased the activity percentage of a synthetic compound library considerably, it was still not as high as in natural-based libraries. Even if the actual 3D structure of the target protein were

used in the creation of a targeted library, it would hardly be more – or even equally – active compared to a nature-originated library, where evolution has selected the molecules that can destroy a variety of microorganism.

8.2 Host cell response to *C. pneumoniae infection* (**IV, V**)

In this part of the study, we used genomics and proteomics in order to define the host cell response to *C. pneumoniae* infection. In gene expression studies, we utilized human microarrays which contained more than 22 000 human genes and measured the changes in the host cell gene expression from the *C. pneumoniae* infected HL cell line at four different time points p.i. The microarray data was used to select new potential drug targets for *C. pneumoniae*, and the most promising target genes were evaluated by silencing their expression with gene-specific siRNA molecules and monitoring the effect of this treatment on the *C. pneumoniae* infection.

In proteomics studies, we compared the protein profile of infected cells to that of uninfected controls by labeling the proteins from the infected and control cells with different label and identified the proteins whose expression was altered because of the infection.

8.2.1 Interpretation of gene expression data (**IV**)

C. pneumoniae infection causes significant changes in host cell gene expression throughout the entire 72 h development cycle. The majority of these changes take place at the early stage of the infection. In the middle part of the cycle, gene expression seems to normalize after *C. pneumoniae* has hidden inside the inclusion in the host cell cytoplasm, but it increases again at the end of the cycle, when the host cell is broken down. The gene expression changes that we observed were due to two different reasons; activation of the host cell defense systems and conscious control of the host cell gene expression by *C. pneumoniae*. When new drug targets are identified, discrimination between these two is paramount, but not necessarily easy. There are genes that are directly linked to the host cell defense systems, and their over-expression is part of the cell defense system, but the majority of genes could theoretically be beneficial or harmful to the infection.

The microarray results from four time points (12 h, 24 h, 48 h, and 72 h) were compared against each other in order to characterize the host cell response at the different stages of the infection. In order to

facilitate the identification of genes related to harmful phenomena, we used the gene ontology (GO) classification (Ashburner *et al.*, 2000). This classification system groups genes in terms of their associations with different biological processes, cellular components, or molecular functions. A biological process describes a biological goal, which is achieved through one or more assemblies of molecular function. Cellular component describes locations at the levels of subcellular structures and macromolecular complexes. Molecular function describes molecular level activities (Harris *et al.*, 2004). A gene can be annotated to any number of ontologies, and annotation to one ontology is independent of its annotation to other ontologies. Currently, there are 5978 different GO terms in the gene ontology database.

The use of GO terms enabled us to discover broader phenomena than just changes in the expression of individual genes. We used a p-value of 0.01 to identify significantly affected GO terms from the GO-classified data. In this case, p-value measured the probability that the changes in gene expression of all genes belonging to a single GO term at one time point, compared to the average gene expression values of corresponding genes at three other time points, is coincidental. **Table 6** contains a collection of the interesting GO terms that were significantly affected by *C. pneumoniae* infection.

With GO annotation, we were able to find phenomena that are already linked to *C. trachomatis* infection, such as Ca^{2+} ion homeostasis (Majeed *et al.*, 1993), glycogen synthesis (Weigent and Jenkin, 1978), Wnt pathway (Prozialeck *et al.*, 2002), hormone receptor activity (Kaushic *et al.*, 2000), prevention of MHC expression (Zhong *et al.*, 2001), and different effects on vesicle transport systems (Hackstadt *et al.*, 1995). In addition, it has been shown in *C. psittaci* that the host cell's energy production is accelerated by infection (Ojcius *et al.*, 1998), and we observed the same phenomenon concerning increased ATP production and glutamate metabolism. Furthermore, some of the phenomena that we discovered using the GO classification have already been reported to be affected by *C. pneumonia* infection, such as the effects on growth factors (Prochnau *et al.*, 2004) and the changes in host cell phosphorylation (Krull *et al.*, 2004). We were also able to find new host cell phenomena not previously known to be affected by *C. pneumonia* infection, such as chaperone regulator activity, amino acid derivative catabolism, and subtilase activity. In addition, we also obtained more detailed information about the potential mechanisms of action of the discovered phenomena, such as Ca^{2+} ion influence through the troponin complex.

Table 6. The most interesting significantly affected GO terms at the different stages of the infection. Complete underlining of the GO term reflects a positive collective change on the genes of that GO term, no underlining reflects a negative change, and underlining of the GO term with a broken line means that about half of the genes of that term were down- and the other half up-regulated. The changes were defined by looking at the gene graphs in the R program.

12 h	24 h	48 h	72 h
Exocyst	Troponin complex	Exocyst	Histone acetyltransferase activity
Clathrin coat of trans-Golgi network vesicle	cGMP-specific phosphodiesterase activity	Subtilase activity	Amino acid transporter activity
v-SNARE activity	Fructose-bisphosphate aldolase activity	Lipid phosphatase activity	Insulin-like growth factor binding
ATPase stimulator activity	Calcium channel regulator activity	Response to toxin	Growth factor binding
Insulin-like growth factor binding	Citrate metabolism	Sterol biosynthesis	Antigen processing, endogenous antigen via MHC class I
Steroid hormone receptor activity	Amino acid derivative catabolism	Xenobiotic metabolism	Negative regulation of Wnt receptor signaling pathway
Chaperone regulator activity	Neurotransmitter catabolism		Lipid catabolism
Sphingomyelin phosphodiesterase activity			
Ligand-dependent nuclear receptor activity			
Negative regulation of Wnt receptor signaling pathway			
Glycogen catabolism			
Protein targeting to vacuole			
DNA modification			

8.2.2 Selection of potential drug target genes (**IV**)

At the first stage, we tried to identify up-regulated GO terms among the significantly affected GO terms at the 12 h time point. These terms were assumed to be essential for *C. pneumoniae* infection, but not for the host cell. GO terms from the 12 h time point were emphasized in the selection process, since prevention of the onset of the infection would eliminate the possible development of chronic infection at

the later stages of the development cycle. Up-regulation was a prerequisite, since we were going to use gene silencing in the target validation process. At the next stage, we identified the individual genes which were mainly responsible for the up-regulation of these GO terms and selected them as target genes. Based on the GO data, we selected a total of five genes as potential targets: CYR61, DKK1, NR4A1, PYG, and V-FOS.

The use of GO terms gives undoubtedly more information about the *C. pneumoniae* – host cell interaction than the expression changes of individual genes. On the other hand, genes that are responsible for the significant effect of a GO term are not necessarily the ones with the largest changes in their expression, when gene expression changes are compared between infected and non-infected cells. The reason for this is that, in the GO classification, the GO terms must be compared against each other at different time points and not against uninfected control. Thus, another set of target genes was selected using gene-wise analysis data, where data from infected cells were compared against control cell data. In this process, individual genes were ranked based on the changes in their expression rate. The genes that showed highest up-regulation by the *C. pneumoniae* infection at the 12 h time point and were potentially important for the survival of the infection, were also selected as target genes. A total of six genes (ACHE, EMP1, EGR1, FLJ23356, FLJ32065, and IGFBP1) were selected as target genes based on the gene-wise analysis.

8.2.3 Drug target validation (**IV**)

The gene expression changes of all selected drug targets were confirmed by using q-PCR (**Table 3 in IV**), and based on the results, the following genes were considered to be validated; CYR61, DKK1, EGR1, IGFBP1, NR4A1, and V-FOS. The validation process showed that the target genes selected based on the GO data were more reliable, since four out of five genes could be validated, whereas only two of the six genes selected using gene-wise analysis showed similar expression in q-PCR and microarray.

The validated genes that were selected using GO data have a variety of functions. The CYR61 gene belongs to the GO term ‘Insulin-like growth factor binding’, and its expression is regulated by the Wnt signaling pathway (Si *et al.*, 2006), which we found to be considerably affected at the different stages of the infection. The DKK1 gene belongs to GO term ’Negative regulation of the Wnt receptor signaling pathway’, and it was selected as a target gene, to further investigate the significance of the Wnt-signaling pathway for *C. pneumoniae* infection. Several steroid-related GO terms were altered

throughout the infection (partially shown in **table 6**), and the NR4A1 gene was selected as a target gene, since it was the most up-regulated gene at the 12 h time point in the 'Steroid hormone receptor activity' GO term. The V-FOS gene was selected as a target gene in order to evaluate the effect of DNA modification on *C. pneumoniae* infection.

Both of the genes that were selected using gene-wise analysis are related to growth factors. The EGR1 gene codes a transcription factor with a C2H2 type zinc finger (Sukhatme *et al.*, 1988). It is normally induced by growth factors and causes expression of the genes required for cell differentiation and mitogenesis. This gene has previously been shown to be activated by *C. pneumoniae* infection (Bea *et al.*, 2003), and it is speculated to be involved in the regulation of proatherosclerotic factors that contribute to the pathogenesis of atherosclerosis (Rupp and Maass, 2003). The IGFBP1 gene is an insulin-like growth factor binding protein (Shimasaki and Ling, 1991), which has potential to modulate the antiapoptotic effect and to regulate cell proliferation (Firth and Baxter, 2002). This family of insulin-like growth factor binding proteins might have an important role in *C. pneumoniae* infection, since the other members of this family are affected during persistent infection (Mannonen *et al.*, 2007).

8.2.4 Effect of target gene silencing on *C. pneumoniae* infection (**IV**)

In order to evaluate the effect of target genes on *C. pneumoniae* infection, two corresponding siRNA molecules were tested for each target, and the silencing effect was monitored with q-PCR (**Table 4 in IV**). The siRNA molecule that lowered gene expression most was used in further experiments. The effect of siRNA molecules on *C. pneumoniae* infection was evaluated using TR-FIA (**Fig. 6**). Each of the target genes was silenced in the HL cell line by transfecting cells with corresponding siRNA and infecting them with *C. pneumoniae* 16h after transfection. The cytotoxic effect of siRNA molecules was also evaluated in order to see whether the silenced genes are suitable for drug targets (**Fig. 6**). In this way we were also able to confirm that the possible reduction in the number of *C. pneumoniae* inclusions was not due to the cytotoxicity of the siRNA molecules used.

All siRNA molecules were able to reduce the number of *C. pneumoniae* particles, but they also caused a small reduction in the number of viable cells. The decreased number of HL cell can explain the smaller number of *C. pneumoniae* particles for some of the siRNAs, such as NR4A1, DKK1, IGFBP1, and possibly V-FOS**,** but not all of them. EGR1 and CYR61 clearly cause the reduction for some other

reasons. These genes are “immediate early genes”, which means that they response fast to an appropriate stimulus.

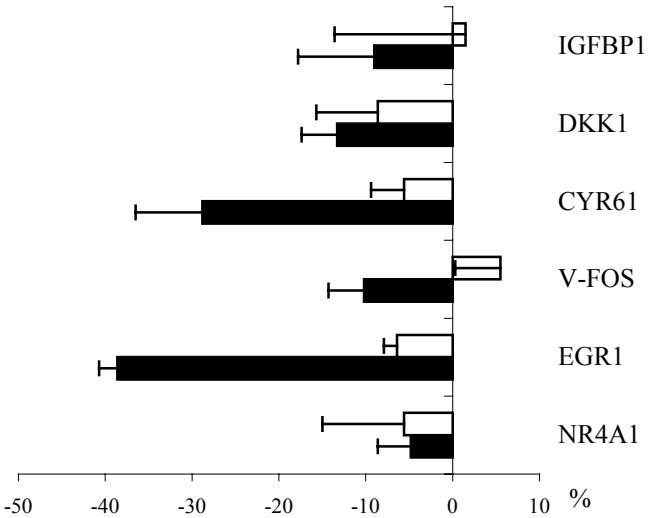

Figure 6. Effect of siRNA molecules on the number of *C. pneumoniae* particles (black bars) and the number of viable HL cells (white bars). Percentage values were calculated from three replicates.

EGR1 encoded a transcription factor is activated by a variety of signals, including serum, growth factors, cytokines, and hormones. It is related to diverse cellular functions, such as cell proliferation, differentiation, and apoptosis (Gashler and Sukhatme, 1995). Based on these experiments, it is impossible to say whether the overall growth-stimulating effect or an as yet unidentified protein whose expression is controlled by this transcription factor makes this gene important for *C. pneumoniae* infection. EGR-1 siRNA silenced the expression of the EGR-1 gene rather poorly (**Table 4 in IV**), which indicates that activation of this gene might be even more important for the survival of *C. pneumoniae* infection than is suggested by these experiments.

The CYR61 encodes a cysteine-rich protein, which is normally induced by growth factors (Lau and Nathans, 1985). This protein is secreted out of the cell, where it binds to heparin and promotes the adhesion of endothelial cells through integrin interaction. It is also capable of promoting angiogenesis and tumor growth (Babic *et al.*, 1998). It is known that *C. pneumonia* (Wuppermann *et al.*, 2001) and *C. trachomatis* (Zhang and Stephens, 1992) attachment to eukaryotic cells is dependent upon the presence

of a heparan sulfate-like ligands on the cell surface, and since the siRNA molecule that silences this gene reduced the number of *C. pneumonia* particles considerably, this protein might promote the attachment of extracellular chlamydia particles to the membranes of surrounding cells. On the other hand, CYR61 is annotated to the GO term 'Insulin-like growth factor binding', and several members of the insulin-like growth factor binding protein family are known to be affected by infection through an unknown mechanism (Lin *et al.*, 2001) (Ren *et al.*, 2003) (Xia *et al.*, 2003). Further studies are therefore needed in order to define the role of CYR61 in chlamydia infection.

8.2.5 Changes at protein level (**V**)

C. pneumoniae infection caused significant changes in the host cell gene expression, and we wanted to see if the same happens at the protein level. Since changes at the protein level take longer time than those at the gene level, we isolated proteins only 48 h and 72 h p.i. We utilized two-dimensional difference gel electrophoresis (2D-DIGE) coupled with mass-spectrometric identification in order to investigate changes in the proteome of infected HL cells. The combination of separation and identification tools that we used enabled identification of 1270 spots at 48 h and 1040 spots at 72 h time points. At both time points, the same four host cell proteins were found to be differently expressed in response to *C. pneumoniae* infection (**Table 2 in V**). All these proteins; Keratin 8 (type II), Keratin 18 (type I), Vimentin, and β-tubulin, are structural components of the cytoskeleton. This seems compatible with the already established finding that rearrangement of the host cell cytoskeleton has an essential role in chlamydia infection (Schramm and Wyrick, 1995) (Wissel *et al.*, 2003).

None of the genes coding for the four proteins whose expression was changed significantly by the *C. pneumoniae* infection showed significant changes in their expression in the microarray experiment. This can be explained by the fact that all of the differently expressed proteins were observed to have different *M*r and/or p*I* values in 2D-gel than they should have in theory. This is normally an indication of post-translational modification, and 2D-DIGE and MALDI-TOF analyses were able to show that these proteins have undergone post-translational modification during the infection. This is not surprising since it is known that tyrosine phosphorylation of host cell proteins takes place at the beginning of *C. trachomatis* infection (Birkelund *et al.*, 1997).

Although gene expression experiments with microchips were able to identify hundreds of genes whose expression was changed dramatically by infection, no change in the corresponding proteins could be

detected in the proteomics experiment. For some proteins, this lack of correlation may be due to the long turnover rate of the protein or some control mechanism that prevents protein translation, but for the most part it is due to the overall lower sensitivity of the protein separation and identification methods (Tian *et al.*, 2004). This lower sensitivity of the proteomics tools can be seen in the number of identified spots. Although there must be tens of thousands of different proteins and their modifications in cells, we were able to separate and identify only about 1000 spots. In addition, only one *C. pneumoniae* protein (the major outer membrane protein; MOMP) could be identified from the 2D-gel.

The proteins that were found to be affected by the *C. pneumoniae* infection in this study were assumed to be unsuitable as drug targets for two reasons: Firstly, they were all products of post-translational modification, and the native forms they were derived from were not affected by the infection, and secondly, all of these proteins are structural elements of the cytoskeleton, and their disturbance would probably have destructive consequences for the host cell.

9 CONCLUSIONS

Natural products are known to exhibit a high level of antibacterial activity. A library of structurally diverse natural compounds (and their derivates) from one of the most active groups, flavonoids, was tested against *C. pneumoniae*. These plant phytoalexins are already known to have a variety of health-promoting effects in humans, and the results of this study revealed one additional effect, namely their extremely high activity against a common human pathogen, *C. pneumonia*. In addition to nature-based compounds, antichlamydial activity was also searched for among synthetic compounds. In order to avoid a massive screening project, synthetic compounds were pre-screened *in silico*. This virtual screening was done using a structural homolog of the selected *C. pneumoniae* target protein. The results of the virtual screening experiment were used to construct a small targeted library. This library was purchased and tested against *C. pneumoniae*, and it demonstrated clearly higher activity than synthetic libraries in general, and although it was not as active as the nature-based library, these compounds provided excellent leads for further synthesis.

The development of a new high-throughput screening system (TR-FIA) for the primary screening of new antichlamydial compounds was one of the aims of this study. The biggest problem in the development of this cell-based assay, the high background value, was resolved by using time-resolved fluorescence label, which can be measured after the background fluorescence has faded away. After the optimization of this partly automated 96-well plate assay, results were reliable, reproducible, and consistent with the two assays mostly used in chlamydia susceptibility testing, IF staining and q-PCR.

Defining the host cell response to *C. pneumoniae* infection and identification of potential drug targets was one of the central purposes of this study. This response was monitored at the gene and protein levels. At the genome level we were able to identify hundreds of genes that were significantly affected by infection. Genes were grouped according to biological processes, cellular components, and molecular functions in order to elicit more information out of the data. This GO classification approach compares the changes in gene expression at the different stages of infection and gives more reliability to the data, since it is not based on the expression changes of a single gene. A group of genes was selected as potential drug target genes based on this classification system, and another group was selected based on gene-wise analysis, where an infected sample is compared against an uninfected control at a single time point. The expression changes of all target genes were confirmed with q-PCR, the validated genes were silenced with corresponding siRNA molecules, and the effect of silencing on *C. pneumonia* infection was monitored using TR-FIA. The more reliable nature of GO data was evident in the validation

process, where four out of five target genes from the GO data could be validated, whereas only two of the six genes that were selected from gene-wise analysis data passed q-PCR validation. The GO classification provided an excellent way to elucidate the *C. pneumonia* infection process, but further research is needed to link these discovered phenomena unquestionably to the *C. pneumonia* infection process. Silencing of the six potential target genes revealed two genes, EGR1 (selected based on gene-wise analysis) and CYR61 (selected based on GO analysis), whose silencing reduced the number of *C. pneumonia* particles considerably without significantly reducing the viability of the host cells.

Although the changes at the protein level are more relevant to cell function, the protein monitoring tools are less sensitive than the tool used to study gene expression, which became evident during this study. We were able to find only four different host cell proteins with significant differences in their expression because of *C. pneumoniae* infection. All these proteins were structural proteins, which supports the earlier observations about the structural rearrangements required for a successful chlamydia infection, but at the same time makes them unsuitable as drug targets.

The main achievements of this study, i.e. the utilization of the GO classification system in drug target identification and the design and construction of small focused libraries with high antichlamydial activity, can also be applied to other bacteria, enabling the development of new antimicrobial compounds against a variety of pathogens.

10 ACKNOWLEDGEMENTS

This work was carried out at the Drug Discovery and Development Technology Center (DDTC), Division of Pharmaceutical Biology, Faculty of Pharmacy, University of Helsinki, during 2002-2007.

I wish to express my gratitude to Professor Raimo Hiltunen, Head of the Division of Pharmaceutical Biology, for his support and for providing excellent working facilities for this thesis.

I am most indebted to my supervisor Professor Pia Vuorela for her practical ideas and advice as well as her continuous positive feedback. I am also deeply grateful for the opportunity to work independently, in order to improve my skills, but still knowing that I could count on her whenever I needed help.

My sincere gratitude goes to Docent Petri Auvinen, who guided and tested me over the years. Although I was one of the visiting scientists, he made me always feel part of his friendly and talented research group.

Warm thanks are due to all my co-authors, especially Professor Heikki Vuorela, Professor Maija Leinonen, Professor Pekka Saikku, Dario Greco, M.Sc., Dr. Kirsi Savijoki, and Docent Nisse Kalkkinen. Without your expertise this work could not have been completed.

Docent Mirja Puolakkainen and Professor Matthias Maas are acknowledged for carefully reviewing the manuscript and for providing valuable comments and suggestions for its improvement.

I am grateful to Sirkka-Liisa Leinonen for revising the language of my publications as well as of the thesis.

Special thanks are due to my roommates; Dr. Tero Wennberg, who shared my interest in the European congress tour and offered decent competition in our moderately scientific competitions over the years, and Dr. Kari Kreander for his entertaining stories and stimulating scientific and personal discussions.

I wish to express my appreciation to Docent Päivi Tammela, who always found time to guide me to misplaced things and to help me in the writing of the numerous reports, applications, and articles, which made this work possible.

I want to thank Anna Galkin, M.Sc., Laura Riihimäki, M.Sc., Päivi Oinonen, M.Sc., and all the other colleagues and the staff at the Division of Pharmaceutical Biology as well as my fellow researchers at DDTC for creating a supportive and pleasant working environment. It truly has been pleasure to work and participate in congresses and seminars with you all.

My dearest thanks go to the people closest to me, my family and friends. My parents, Simo and Sinikka Alvesalo, who have motivated and financially supported me through my studies. I am most fortunate to have my encouraging parents and my sister Sesilia and brother Lari with their families around me. Finally, my most sincere thanks go to my wife Hanne and our three wonderful children Leevi, Lotta, and Linda. You all created a loving atmosphere where I could relax and forget the troubles at work. It truly meant a world to me.

I also wish to express my gratitude to the National Technology Agency of Finland (grants 40708/00 and 40201/04), The Finnish Cultural Foundation (Elli Turunen Foundation), the Finnish Pharmaceutical Society, and the FOTY ry for their financial contributions during this study.

Helsinki, March 2007

11 REFERENCES

Andraws R, Berger JS, Brown DL. (2005). Effects of antibiotic therapy on outcomes of patients with coronary artery disease: a meta-analysis of randomized controlled trials. *JAMA* 293(21):2641-2647.

Apfalter P, Blasi F, Boman J, Gaydos CA, Kundi M, Maass M, *et al*. (2001). Multicenter comparison trial of DNA extraction methods and PCR assays for detection of *Chlamydia pneumoniae* in endarterectomy specimens. *J Clin Microbiol* 39(2):519-524.

Arthur JM. (2003). Proteomics. *Curr Opin Nephrol Hypertens* 12(4):423-430.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25-29.

Babic AM, Kireeva ML, Kolesnikova TV, Lau LF. (1998). CYR61, a product of a growth factor-inducible immediate early gene, promotes angiogenesis and tumor growth. *Proc Natl Acad Sci U S A* 95(11):6355-6360.

Baltch AL, Smith RP, Ritz WJ, Carpenter AN, Bopp LH, Michelsen PB, *et al*. (2004). Effect of levofloxacin on the viability of intracellular *Chlamydia pneumoniae* and modulation of proinflammatory cytokine production by human monocytes. *Diagn Microbiol Infect Dis* 50(3):205-212.

Balunas MJ, Kinghorn AD. (2005). Drug discovery from medicinal plants. *Life Sci* 78(5):431-441.

Bea F, Puolakkainen MH, McMillen T, Hudson FN, Mackman N, Kuo CC, *et al*. (2003). *Chlamydia pneumoniae* induces tissue factor expression in mouse macrophages via activation of Egr-1 and the MEK-ERK1/2 pathway. *Circ Res* 92(4):394-401.

Beatty WL, Byrne GI, Morrison RP. (1993). Morphologic and antigenic characterization of interferon gamma-mediated persistent *Chlamydia trachomatis* infection in vitro. *Proc Natl Acad Sci U S A* 90(9):3998-4002.

Birkelund S, Bini L, Pallini V, Sanchez-Campillo M, Liberatori S, Clausen JD, *et al*. (1997). Characterization of *Chlamydia trachomatis* l2-induced tyrosine-phosphorylated HeLa cell proteins by two-dimensional gel electrophoresis. *Electrophoresis* 18(3-4):563-567.

Butler MS, Buss AD. (2006). Natural products--the future scaffolds for novel antibiotics? *Biochem Pharmacol* 71(7):919-929.

C. elegans Sequencing Consortium. (1998). Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282(5396):2012-2018.

Caldwell HD, Kromhout J, Schachter J. (1981). Purification and partial characterization of the major outer membrane protein of *Chlamydia trachomatis*. *Infect Immun* 31(3):1161-1176.

Carey KL, Westwood NJ, Mitchison TJ, Ward GE. (2004). A small-molecule approach to studying invasive mechanisms of *Toxoplasma gondii*. *Proc Natl Acad Sci U S A* 101(19):7433-7438.

Chi EY, Kuo CC, Grayston JT. (1987). Unique ultrastructure in the elementary body of Chlamydia sp. strain TWAR. *J Bacteriol* 169(8):3757-3763.

Chopra I, Storey C, Falla TJ, Pearce JH. (1998). Antibiotics, peptidoglycan synthesis and genomics: the chlamydial anomaly revisited. *Microbiology* 144 (Pt 10)(Pt 10):2673-2678.

Cles LD, Stamm WE. (1990). Use of HL cells for improved isolation and passage of *Chlamydia pneumoniae*. *J Clin Microbiol* 28(5):938-940.

Coombes BK, Mahony JB. (2001). cDNA array analysis of altered gene expression in human endothelial cells in response to *Chlamydia pneumoniae* infection. *Infect Immun* 69(3):1420-1427.

Cragg GM, Newman DJ. (2005). Plants as a source of anti-cancer agents. *J Ethnopharmacol* 100(1-2):72-79.

Danesh J. (2005). Antibiotics in the prevention of heart attacks. *Lancet* 365(9457):365-367.

DiMasi JA, Hansen RW, Grabowski HG. (2003). The price of innovation: new estimates of drug development costs. *J Health Econ* 22(2):151-185.

Ekman MR, Grayston JT, Visakorpi R, Kleemola M, Kuo CC, Saikku P. (1993a). An epidemic of infections due to *Chlamydia pneumoniae* in military conscripts. *Clin Infect Dis* 17(3):420-425.

Epstein SE, Zhou YF, Zhu J. (1999). Infection and atherosclerosis: emerging mechanistic paradigms. *Circulation* 100(4):e20-8.

Feher M, Schmidt JM. (2003). Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J Chem Inf Comput Sci* 43(1):218-227.

Firth SM, Baxter RC. (2002). Cellular actions of the insulin-like growth factor binding proteins. *Endocr Rev* 23(6):824-854.

Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, *et al*. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* 269(5223):496-512.

Furka A, Sebestyen F, Asgedom M, Dibo G. (1991). General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res* 37(6):487-493.

Gashler A, Sukhatme VP. (1995). Early growth response protein 1 (Egr-1): prototype of a zinc-finger family of transcription factors. *Prog Nucleic Acid Res Mol Biol* 50:191-224.

Gerhold DL, Jensen RV, Gullans SR. (2002). Better therapeutics through microarrays. *Nat Genet* 32 Suppl:547-551.

Gershell LJ, Atkins JH. (2003). A brief history of novel drug discovery technologies. *Nat Rev Drug Discov* 2(4):321-327.

Gieffers J, Fullgraf H, Jahn J, Klinger M, Dalhoff K, Katus HA*, et al*. (2001). *Chlamydia pneumoniae* infection in circulating human monocytes is refractory to antibiotic treatment. *Circulation* 103(3):351-356.

Graul AI, Prous JR. (2006). The year's new drugs. *Drug News Perspect* 19(1):33-53.

Grayston JT, Kuo CC, Campbell LA, Wang SP. (1989). *Chlamydia pneumoniae* sp. nov. for *Chlamydia* sp. strain TWAR. *Int J Syst Bacteriol 39:88-90.*

Grayston JT. (2000). Background and current knowledge of *Chlamydia pneumoniae* and atherosclerosis. *J Infect Dis* 181 Suppl 3:S402-10.

Grayston JT. (1992). Infections caused by *Chlamydia pneumoniae* strain TWAR. *Clin Infect Dis* 15(5):757-761.

Grayston JT, Kuo CC, Wang SP, Altman J. (1986). A new *Chlamydia psittaci* strain, TWAR, isolated in acute respiratory tract infections. *N Engl J Med* 315(3):161-168.

Grayston JT, Campbell LA, Kuo CC, Mordhorst CH, Saikku P, Thom DH*, et al*. (1990). A new respiratory tract pathogen: *Chlamydia pneumoniae* strain TWAR. *J Infect Dis* 161(4):618-625.

Hackstadt T, Scidmore MA, Rockey DD. (1995). Lipid metabolism in *Chlamydia trachomatis*-infected cells: directed trafficking of Golgi-derived sphingolipids to the chlamydial inclusion. *Proc Natl Acad Sci U S A* 92(11):4877-4881.

Hahn DL. (1999). *Chlamydia pneumoniae*, asthma, and COPD: what is the evidence? *Ann Allergy Asthma Immunol* 83(4):271-88, 291; quiz 291-2.

Hammerschlag MR, Hyman CL, Roblin PM. (1992a). In vitro activities of five quinolones against *Chlamydia pneumoniae*. *Antimicrob Agents Chemother* 36(3):682-683.

Hammerschlag MR, Chirgwin K, Roblin PM, Gelling M, Dumornay W, Mandel L, *et al*. (1992b). Persistent infection with *Chlamydia pneumoniae* following acute respiratory illness. *Clin Infect Dis* 14(1):178-182.

Hanash S. (2003). Disease proteomics. *Nature* 422(6928):226-232.

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, *et al*. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32(Database issue):D258-61.

Hatch TP, Allan I, Pearce JH. (1984). Structural and polypeptide differences between envelopes of infective and reproductive life cycle forms of Chlamydia spp. *J Bacteriol* 157(1):13-20.

Hess S, Peters J, Bartling G, Rheinheimer C, Hegde P, Magid-Slav M, *et al*. (2003). More than just innate immunity: comparative analysis of *Chlamydophila pneumoniae* and *Chlamydia trachomatis* effects on host-cell gene regulation. *Cell Microbiol* 5(11):785-795.

Hogan RJ, Mathews SA, Mukhopadhyay S, Summersgill JT, Timms P. (2004). Chlamydial persistence: beyond the biphasic paradigm. *Infect Immun* 72(4):1843-1855.

Janssen P, Audit B, Cases I, Darzentas N, Goldovsky L, Kunin V, *et al*. (2003). Beyond 100 genomes. *Genome Biol* 4(5):402.

Jokinen C, Heiskanen L, Juvonen H, Kallinen S, Kleemola M, Koskela M, *et al*. (2001). Microbial etiology of community-acquired pneumonia in the adult population of 4 municipalities in eastern Finland. *Clin Infect Dis* 32(8):1141-1154.

Kalman S, Mitchell W, Marathe R, Lammel C, Fan J, Hyman RW, *et al*. (1999). Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nat Genet* 21(4):385-389.

Kaukoranta-Tolvanen SS, Laitinen K, Saikku P, Leinonen M. (1994). *Chlamydia pneumoniae* multiplies in human endothelial cells in vitro. *Microb Pathog* 16(4):313-319.

Kaushic C, Zhou F, Murdin AD, Wira CR. (2000). Effects of estradiol and progesterone on susceptibility and early immune responses to *Chlamydia trachomatis* infection in the female reproductive tract. *Infect Immun* 68(7):4207-4216.

Klose J, Kobalz U. (1995). Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome. *Electrophoresis* 16(6):1034-1059.

Koehler L, Nettelnbreker E, Hudson AP, Ott N, Gerard HC, Branigan PJ, *et al*. (1997). Ultrastructural and molecular analyses of the persistence of *Chlamydia trachomatis* (serovar K) in human monocytes. *Microb Pathog* 22(3):133-142.

Koehn FE, Carter GT. (2005). The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* 4(3):206-220.

Krull M, Kramp J, Petrov T, Klucken AC, Hocke AC, Walter C, *et al*. (2004). Differences in cell activation by *Chlamydophila pneumoniae* and *Chlamydia trachomatis* infection in human endothelial cells. *Infect Immun* 72(11):6615-6621.

Kubo I, Fujita K, Nihei K, Masuoka N. (2003). Non-antibiotic antibacterial activity of dodecyl gallate. *Bioorg Med Chem* 11(4):573-580.

Kuo CC, Grayston JT. (1988). In vitro drug susceptibility of Chlamydia sp. strain TWAR. *Antimicrob Agents Chemother* 32(2):257-258.

Kuo CC, Jackson LA, Campbell LA, Grayston JT. (1995). *Chlamydia pneumoniae* (TWAR). *Clin Microbiol Rev* 8(4):451-461.

Kutlin A, Roblin PM, Hammerschlag MR. (1999). In vitro activities of azithromycin and ofloxacin against *Chlamydia pneumoniae* in a continuous-infection model. *Antimicrob Agents Chemother* 43(9):2268-2272.

Kutlin A, Roblin PM, Hammerschlag MR. (2002). Effect of prolonged treatment with azithromycin, clarithromycin, or levofloxacin on *Chlamydia pneumoniae* in a continuous-infection Model. *Antimicrob Agents Chemother* 46(2):409-412.

Lau LF, Nathans D. (1985). Identification of a set of genes expressed during the G0/G1 transition of cultured mouse cells. *EMBO J* 4(12):3145-3151.

Lee ML, Schneider G. (2001). Scaffold architecture and pharmacophoric properties of natural products and trade drugs: application in the design of natural product-based combinatorial libraries. *J Comb Chem* 3(3):284-289.

Leinonen M, Saikku P. (2002). Evidence for infectious agents in cardiovascular disease and atherosclerosis. *Lancet Infect Dis* 2(1):11-17.

Lewis LM, Engle LJ, Pierceall WE, Hughes DE, Shaw KJ. (2004). Affinity capillary electrophoresis for the screening of novel antimicrobial targets. *J Biomol Screen* 9(4):303-308.

Libby P, Egan D, Skarlatos S. (1997). Roles of infectious agents in atherosclerosis and restenosis: an assessment of the evidence and need for future research. *Circulation* 96(11):4095-4103.

Lin TM, Campbell LA, Rosenfeld ME, Kuo CC. (2001). Human monocyte-derived insulin-like growth factor-2 enhances the infection of human arterial endothelial cells by *Chlamydia pneumoniae*. *J Infect Dis* 183(9):1368-1372.

Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Ad*V *Drug Deli*V*ery Re*V *23*, 3-25.

Liu B, Li S, Hu J. (2004). Technological advances in high-throughput screening. *Am J Pharmacogenomics* 4(4):263-276.

Lugert R, Kuhns M, Polch T, Gross U. (2004). Expression and localization of type III secretion-related proteins of C*hlamydia pneumoniae*. *Med Microbiol Immunol (Berl)* 193(4):163-171.

Macko V, Stegemann H. (1969). Mapping of potato proteins by combined electrofocusing and electrophoresis identification of varieties. *Hoppe Seylers Z Physiol Chem* 350(7):917-919.

Maillard JY. (2002). Bacterial target sites for biocide action. *J Appl Microbiol* 92 Suppl:16S-27S.

Majeed M, Gustafsson M, Kihlstrom E, Stendahl O. (1993). Roles of Ca2+ and F-actin in intracellular aggregation of *Chlamydia trachomatis* in eucaryotic cells. *Infect Immun* 61(4):1406-1414.

Mannonen L, Nikula T, Haveri A, Reinikainen A, Vuola JM, Lahesmaa R*, et al*. (2007). Up-regulation of host cell genes during interferon-gamma-induced persistent *Chlamydia pneumoniae* infection in HL cells. *J Infect Dis* 195(2):212-219.

Manthey JA. (2000). Biological properties of flavonoids pertaining to inflammation. *Microcirculation* 7(6 Pt 2):S29-34.

Mathews SA, Timms P. (2000). Identification and mapping of sigma-54 promoters in *Chlamydia trachomatis. J Bacteriol* 182(21):6239-6242.

Matsumoto A, Manire GP. (1970). Electron microscopic observations on the effects of penicillin on the morphology of *Chlamydia psittaci*. *J Bacteriol* 101(1):278-285.

McCoy AJ, Sandlin RC, Maurelli AT. (2003). In vitro and in vivo functional activity of Chlamydia MurA, a UDP-N-acetylglucosamine enolpyruvyl transferase involved in peptidoglycan synthesis and fosfomycin resistance. *J Bacteriol* 185(4):1218-1228.

Middleton E,Jr, Kandaswami C. (1992). Effects of flavonoids on immune and inflammatory cell functions. *Biochem Pharmacol* 43(6):1167-1179.

Miyashita N, Niki Y, Kishimoto T, Nakajima M, Matsushima T. (1997). In vitro and in vivo activities of AM-1155, a new fluoroquinolone, against Chlamydia spp. *Antimicrob Agents Chemother* 41(6):1331-1334.

Moazed TC, Kuo CC, Grayston JT, Campbell LA. (1998). Evidence of systemic dissemination of *Chlamydia pneumoniae* via macrophages in the mouse. *J Infect Dis* 177(5):1322-1325.

Moulder JW. (1966). The relation of the psittacosis group (Chlamydiae) to bacteria and viruses. *Annu Rev Microbiol* 20:107-130.

Nathan C. (2004). Antibiotics at the crossroads. *Nature* 431(7011):899-902.

Newhall WJ, Batteiger B, Jones RB. (1982). Analysis of the human serological response to proteins of *Chlamydia trachomatis*. *Infect Immun* 38(3):1181-1189.

Newman DJ, Cragg GM, Snader KM. (2003). Natural products as sources of new drugs over the period 1981-2002. *J Nat Prod* 66(7):1022-1037.

Ojala T, Remes S, Haansuu P, Vuorela H, Hiltunen R, Haahtela K, *et al*. (2000). Antimicrobial activity of some coumarin containing herbal plants growing in Finland. *J Ethnopharmacol* 73(1-2):299-305.

Ojcius DM, Degani H, Mispelter J, Dautry-Varsat A. (1998). Enhancement of ATP levels and glucose metabolism during an infection by Chlamydia. NMR studies of living cells. *J Biol Chem* 273(12):7052-7058.

Olson JA,Jr. (2004). Application of microarray profiling to clinical trials in cancer. *Surgery* 136(3):519-523.

Onyango P. (2004). The role of emerging genomics and proteomics technologies in cancer drug target discovery. *Curr Cancer Drug Targets* 4(2):111-124.

Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. (2007). Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* 6(1):29-40.

Pedersen LB, Birkelund S, Christiansen G. (1994). Interaction of the *Chlamydia trachomatis* histone H1-like protein (Hc1) with DNA and RNA causes repression of transcription and translation in vitro. *Mol Microbiol* 11(6):1085-1098.

Pelaez F. (2006). The historical delivery of antibiotics from microbial natural products--can history repeat? *Biochem Pharmacol* 71(7):981-990.

Persidis A. (1998). Proteomics. *Nat Biotechnol* 16(4):393-394.

Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, *et al.* (2002). Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 359(9306):572-577.

Pilger BD, Cui C, Coen DM. (2004). Identification of a small molecule that inhibits herpes simplex virus DNA Polymerase subunit interactions and viral replication. *Chem Biol* 11(5):647-654.

Prochnau D, Rodel J, Hartmann M, Straube E, Figulla HR. (2004). Growth factor production in human endothelial cells after *Chlamydia pneumoniae* infection. *Int J Med Microbiol* 294(1):53-57.

Prozialeck WC, Fay MJ, Lamar PC, Pearson CA, Sigar I, Ramsey KH. (2002). *Chlamydia trachomatis* disrupts N-cadherin-dependent cell-cell junctions and sequesters beta-catenin in human cervical epithelial cells. *Infect Immun* 70(5):2605-2613.

Puolakkainen M, Kuo CC, Campbell LA. (2005). *Chlamydia pneumoniae* uses the mannose 6-phosphate/insulin-like growth factor 2 receptor for infection of endothelial cells. *Infect Immun* 73(8):4620-4625.

Pusztai L, Ayers M, Stec J, Hortobagyi GN. (2003). Clinical application of cDNA microarrays in oncology. *Oncologist* 8(3):252-258.

Puupponen-Pimiä R, Nohynek L, Alakomi HL, Oksman-Caldentey KM. (2005). Bioactive berry compounds-novel tools against human pathogens. *Appl Microbiol Biotechnol* 67(1):8-18.

Rauha JP, Remes S, Heinonen M, Hopia A, Kähkonen M, Kujala T, *et al*. (2000). Antimicrobial effects of Finnish plant extracts containing flavonoids and other phenolic compounds. *Int J Food Microbiol* 56(1):3-12.

Raulston JE. (1997). Response of *Chlamydia trachomatis* serovar E to iron restriction in vitro and evidence for iron-regulated chlamydial proteins. *Infect Immun* 65(11):4539-4547.

Ren Q, Robertson SJ, Howe D, Barrows LF, Heinzen RA. (2003). Comparative DNA microarray analysis of host cell transcriptional responses to infection by *Coxiella burnetii* or *Chlamydia trachomatis*. *Ann N Y Acad Sci* 990:701-713.

Roblin PM, Hammerschlag MR. (1998). In vitro activity of a new ketolide antibiotic, HMR 3647, against *Chlamydia pneumoniae*. *Antimicrob Agents Chemother* 42(6):1515-1516.

Rockey DD, Heinzen RA, Hackstadt T. (1995). Cloning and characterization of a *Chlamydia psittaci* gene coding for a protein localized in the inclusion membrane of infected cells. *Mol Microbiol* 15(4):617-626.

Rockey DD, Grosenbach D, Hruby DE, Peacock MG, Heinzen RA, Hackstadt T. (1997). *Chlamydia psittaci* IncA is phosphorylated by the host cell and is exposed on the cytoplasmic face of the developing inclusion. *Mol Microbiol* 24(1):217-228.

Rupp J, Maass M. (2003). Egr-1, a major link between infection and atherosclerosis? *Circ Res* 92(9):e78.

Saikku P, Wang SP, Kleemola M, Brander E, Rusanen E, Grayston JT. (1985). An epidemic of mild pneumonia due to an unusual strain of Chlamydia psittaci. *J Infect Dis* 151(5):832-839.

Saikku P, Leinonen M, Mattila K, Ekman MR, Nieminen MS, Mäkelä PH, *et al*. (1988). Serological evidence of an association of a novel Chlamydia, TWAR, with chronic coronary heart disease and acute myocardial infarction. *Lancet* 2(8618):983-986.

Schapira M, Abagyan R, Totrov M. (2003). Nuclear hormone receptor targeted virtual screening. *J Med Chem* 46(14):3045-3059.

Schena M, Shalon D, Davis RW, Brown PO. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467-470.

Schmid EF, Smith DA. (2004). Is pharmaceutical R&D just a game of chance or can strategy make a difference? *Drug Discov Today* 9(1):18-26.

Schramm N, Wyrick PB. (1995). Cytoskeletal requirements in *Chlamydia trachomatis* infection of host cells. *Infect Immun* 63(1):324-332.

Schubert CM. (2003). Microarray to be used as routine clinical screen. *Nat Med* 9(1):9.

Seppälä H, Skurnik M, Soini H, Roberts MC, Huovinen P. (1998). A novel erythromycin resistance methylase gene (*erm*TR) in Streptococcus pyogenes. *Antimicrob Agents Chemother* 42(2):257-262.

Shaw AC, Vandahl BB, Larsen MR, Roepstorff P, Gevaert K, Vandekerckhove J, *et al*. (2002). Characterization of a secreted Chlamydia protease. *Cell Microbiol* 4(7):411-424.

Shi Y, Tokunaga O. (2004). *Chlamydia pneumoniae* (*C. pneumoniae*) infection upregulates atherosclerosis-related gene expression in human umbilical vein endothelial cells (HUVECs). *Atherosclerosis* 177(2):245-253.

Shimasaki S, Ling N. (1991). Identification and molecular characterization of insulin-like growth factor binding proteins (IGFBP-1, -2, -3, -4, -5 and -6). *Prog Growth Factor Res* 3(4):243-266.

Si W, Kang Q, Luu HH, Park JK, Luo Q, Song WX, *et al*. (2006). CCN1/Cyr61 is regulated by the canonical Wnt signal and plays an important role in Wnt3A-induced osteoblast differentiation of mesenchymal stem cells. *Mol Cell Biol* 26(8):2955-2964.

Singh SB, Barrett JF. (2006). Empirical antibacterial drug discovery--foundation in natural products. *Biochem Pharmacol* 71(7):1006-1015.

Soini E, Kojola H. (1983). Time-resolved fluorometer for lanthanide chelates--a new generation of nonisotopic immunoassays. *Clin Chem* 29(1):65-68.

Spaller MR, Burger MT, Fardis M, Bartlett PA. (1997). Synthetic strategies in combinatorial chemistry. *Curr Opin Chem Biol* 1(1):47-53.

Stahura FL, Godden JW, Xue L, Bajorath J. (2000). Distinguishing between natural products and synthetic molecules by descriptor Shannon entropy analysis and binary QSAR calculations. *J Chem Inf Comput Sci* 40(5):1245-1252.

Strohl WR. (2000). The role of natural products in a modern drug discovery program. *Drug Discov Today* 5(2):39-41.

Sukhatme VP, Cao XM, Chang LC, Tsai-Morris CH, Stamenkovich D, Ferreira PC, *et al*. (1988). A zinc finger-encoding gene coregulated with c-fos during growth and differentiation, and after cellular depolarization. *Cell* 53(1):37-43.

Tammela P, Laitinen L, Galkin A, Wennberg T, Heczko R, Vuorela H, *et al*. (2004). Permeability characteristics and membrane affinity of flavonoids and alkyl gallates in Caco-2 cells and in phospholipid vesicles. *Arch Biochem Biophys* 425(2):193-199.

Tian Q, Stepaniants SB, Mao M, Weng L, Feetham MC, Doyle MJ, *et al*. (2004). Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Mol Cell Proteomics* 3(10):960-969.

Tondella ML, Talkington DF, Holloway BP, Dowell SF, Cowley K, Soriano-Gabarro M, *et al*. (2002). Development and evaluation of real-time PCR-based fluorescence assays for detection of C*hlamydia pneumoniae*. *J Clin Microbiol* 40(2):575-583.

Törmakangas L, Vuorela P, Saario E, Leinonen M, Saikku P, Vuorela H. (2005). In vivo treatment of acute *Chlamydia pneumoniae* infection with the flavonoids quercetin and luteolin and an alkyl gallate, octyl gallate, in a mouse model. *Biochem Pharmacol* 70(8):1222-1230.

Van Eyk JE. (2001). Proteomics: unraveling the complexity of heart disease and striving to change cardiology. *Curr Opin Mol Ther* 3(6):546-553.

Vandahl BB, Stensballe A, Roepstorff P, Christiansen G, Birkelund S. (2005). Secretion of Cpn0796 from *Chlamydia pneumoniae* into the host cell cytoplasm by an autotransporter mechanism. *Cell Microbiol* 7(6):825-836.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, *et al*. (2001). The sequence of the human genome. *Science* 291(5507):1304-1351.

Virok D, Loboda A, Kari L, Nebozhyn M, Chang C, Nichols C, *et al*. (2003). Infection of U937 monocytic cells with *Chlamydia pneumoniae* induces extensive changes in host cell gene expression. *J Infect Dis* 188(9):1310-1321.

Wang SP, Grayston JT. (1970). Immunologic relationship between genital TRIC, lymphogranuloma venereum, and related organisms in a new microtiter indirect immunofluorescence test. *Am J Ophthalmol* 70(3):367-374.

Weeraratna AT, Nagel JE, de Mello-Coelho V, Taub DD. (2004). Gene expression profiling: from microarrays to medicine. *J Clin Immunol* 24(3):213-224.

Weigent DA, Jenkin HM. (1978). Contrast of Glycogenesis and protein synthesis in monkey kidney cells and HeLa cells infected with *Chlamydia trachomatis* lymphogranuloma venereum. *Infect Immun* 20(3):632-639.

Wetmore BA, Merrick BA. (2004). Toxicoproteomics: proteomics applied to toxicology and pathology. *Toxicol Pathol* 32(6):619-642.

Wissel H, Schulz C, Rudiger M, Krull M, Stevens PA, Wauer RR. (2003). *Chlamydia pneumoniae* affect surfactant trafficking and secretion due to changes of type II cell cytoskeleton. *Am J Respir Cell Mol Biol* 29(3 Pt 1):303-313.

Wolf K, Hackstadt T. (2001). Sphingomyelin trafficking in *Chlamydia pneumoniae*-infected cells. *Cell Microbiol* 3(3):145-152.

Wolf K, Fischer E, Hackstadt T. (2000). Ultrastructural analysis of developmental events in *Chlamydia pneumoniae*-infected cells. *Infect Immun* 68(4):2379-2385.

Wuppermann FN, Hegemann JH, Jantos CA. (2001). Heparan sulfate-like glycosaminoglycan is a cellular receptor for Chlamydia pneumoniae. *J Infect Dis* 184(2):181-187.

Xia M, Bumgarner RE, Lampe MF, Stamm WE. (2003). *Chlamydia trachomatis* infection alters host cell transcription in diverse cellular pathways. *J Infect Dis* 187(3):424-434.

Zhang JH, Chung TD, Oldenburg KR. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* 4(2):67-73.

Zhang JP, Stephens RS. (1992). Mechanism of *C. trachomatis* attachment to eukaryotic host cells. *Cell* 69(5):861-869.

Zhong G, Fan P, Ji H, Dong F, Huang Y. (2001). Identification of a chlamydial protease-like activity factor responsible for the degradation of host transcription factors. *J Exp Med* 193(8):935-942.

Zolli-Juran M, Cechetto JD, Hartlen R, Daigle DM, Brown ED. (2003). High throughput screening identifies novel inhibitors of Escherichia coli dihydrofolate reductase that are competitive with dihydrofolate. *Bioorg Med Chem Lett* 13(15):2493-2496.